A TEXT-BOOK OF MATHEMATICAL ANALYSIS

THE UNIFORM CALCULUS AND ITS APPLICATIONS

BY

R. L. GOODSTEIN

OXFORD
AT THE CLARENDON PRESS

Oxford University Press
BOMBAY CALCUTTA MADRAS

FIRST EDITION 1948
REPRINTED 1949

PRINTED IN GREAT BRITAIN

PREFACE

Text-books on the Differential and Integral Calculus for schoolboys and university students fall into two main classes; those which are based, implicitly or explicitly, on geometrical intuition, and which overcome the difficulties in an elementary presentation of the fundamental theorems either by omission or falsification, and those which develop the subject rigorously on the basis of the Dedekind number theory. The Dedekind theory, or an equivalent formulation, is suitable only for the mathematical specialist, and the student who seeks to acquire a knowledge of the Calculus for application in other fields is obliged to depend upon an account of the subject which leaves him the servant, and never the master, of a fundamental technique.

After considering this dilemma for many years, and after many fruitless attempts to reconcile the claims of these two methods, I came to the conclusion that the solution of the problem lay in a simplification of the Calculus itself.

The simplified system, it is clear, must fulfil two principal requirements; firstly, it must have the full technical power of the current system, i.e. must serve to find rates of change, areas, centres of gravity, and generally must in all its applications differ in no respect from the classical Calculus. All 'ordinary' functions which are differentiable or integrable in the current system must be differentiable or integrable in the simplified system (and of course the results of these operations must be unchanged). Secondly, the system must be founded in ideas so simple that they can be appreciated by any student with a School Certificate knowledge of arithmetic and algebra.

It is believed that the 'Uniform Calculus', in the main, fulfils both these requirements. No doubt, time and experience will bring to light modifications, or even drastically different methods of development, which accomplish the task far better; no one looks forward to that progress more eagerly than the present author.

The less the Uniform Calculus is found to diverge from the classical, in practice, the better will one of its objects have been

accomplished. The student who already has some experience in ordinary differentiation and integration will be able to use this work without perceiving any difference at all; the following observations are directed, not to him, but to the expert.

One of the more difficult, but essential, steps in the current development of the Calculus is the proof of the uniform continuity of a function which is continuous in a closed interval. The difficulty is of course obviated by taking uniform continuity as the *primary* attribute, and this procedure is practicable since the familiar proofs of continuity of the elementary functions serve equally to prove uniform continuity. So, too, the simplest proofs of a surprisingly large number of elementary theorems, e.g. the fundamental theorem $\int_a^x f'(t)\,dt = f(x) - f(a)$, and theorems on the inversion of the order of repeated differentiations, require a continuous derivative, and it seems preferable to obtain this continuity as a consequence of *uniform differentiability*, rather than to postulate continuity *ad hoc*. The proofs of the uniform differentiability of the elementary functions are no more difficult than the proofs of their differentiability at a point.

It is of course true that there are functions differentiable in the ordinary sense but not uniformly differentiable, but these may fairly be described as uncommon functions with which an elementary work is not concerned. A well-known instance of such a function is $f(x)$ defined by the conditions

$$f(0) = 0, \qquad f(x) = x^2\sin(1/x), \quad x \neq 0;$$

this function is everywhere differentiable in the classical sense, with a derivative discontinuous at the origin, and therefore it is not uniformly differentiable in any interval which contains the origin.

The 'Uniform Calculus' is not, however, a mere collection of those parts of the current theory which treat of uniform processes; for one thing what has hitherto been considered is far from forming a comprehensive system, and—what is more important—there is no indication in the current theory that a calculus of uniform processes is susceptible of a more elementary *foundation* than the calculus of non-uniform processes.

The restriction to uniformity (and other simplifications) make

it possible, within the field covered by this work, to dispense with such notions as sections and exact bounds, and with the classical foundation theorems like that of Heine-Borel on covering intervals, and the Weierstrass theorem on the existence of a limit-point of a bounded set (or its weaker form which asserts the convergence of a monotonic bounded sequence).

We have taken as our foundation-stone the notion of a convergent sequence of terminating decimals. Such a sequence is shown to determine, digit by digit, an endless decimal which is called the limit of the sequence. In the first chapter this idea is developed in a purely practical manner, emphasis being laid on the numerical determination of the limit in actual cases; the formal proof of the existence of the limit is given in the Appendix. The next step is to show that no new difficulty arises when the terms of the convergent sequence are themselves endless decimals, and this result is equivalent to Dedekind's theorem. A somewhat similar treatment of convergence is to be found in Baire's famous paper 'Fonctions discontinues', *Coll. Borel*, 1905, and *Acta Math.*, 1908, though the order of procedure is reversed and our defining properties are there derived from the classical formulation.

The standard functions like e^x, $\sin x$, $\cos x$ are *defined* by their expansions. The power series is a natural extension of the polynomial, and the step from polynomial to power series presents no difficulty to the student who has seen the development of the number concept from terminating to endless decimal. The simplest power series from this point of view are the geometric series, and the series which can be derived from it by differentiation and integration. The exponential series is then introduced as the series which is unchanged by differentiation; taking the odd and even powers of the exponential series separately we meet the hyperbolic functions, and then, alternating the signs, we reach the circular functions. By these steps the fundamental series arise in a natural manner and the beginner is not left to wonder why just these series are studied in detail; the periodicity of the circular functions must however appear as a remarkable and unexpected consequence of the addition theorems. The only natural definitions of the circular functions which make the periodicity immediately apparent (apart from the trivial one in which an inverse circular

function is defined for a certain range, as for instance by a definite integral, and then *extended* beyond this range by postulating periodicity) are either by the expansion in partial fractions

$$\frac{1}{\sin x} = \frac{1}{x} + \sum_{n=1}^{\infty} (-1)^n \frac{2x}{x^2 - n^2\pi^2},$$

or by the expression of $\sin x$ as an endless product $x \prod_{n=1}^{\infty} (1 - x^2/n^2\pi^2)$, but these definitions do not lend themselves to a simple development.

One point of notation requires special mention. The notion of a decimal zero to n places of decimals, i.e. a number of absolute value less than $1/10^n$, is required so frequently that a specific notation is used for it; we have denoted by $0(n)$ any decimal which commences with a sequence of n zeros, and have expressed the fact that a particular decimal x is of this form by the equation $x = 0(n)$. This notation was reluctantly adopted only after long and careful consideration, its main drawback being of course that the similar sign '$o(n)$' has already a well-established usage in the theory of orders of magnitude. This theory has no part in the present work so there is no immediate danger of confusion, but it would certainly have been preferable to avoid it if possible. Unfortunately no other notation, invented for the purpose, served as well or suggested the desired idea so forcibly as the one adopted. The abbreviation '$\phi(n) = 0(r)$, $n \geqslant n_r$', standing for '$\phi(n) = 0(r)$ for *any* n not less than n_r' (the term 'any' has the universal not the existential connotation, throughout the book) is used extensively and is introduced quite early in the first chapter.

There are no diagrams in this book. All the propositions and proof processes are purely algebraic, even though the language of geometry is occasionally used as a more vivid form of expression, for instance when we talk of a number between a and b as a *point* in the interval (a, b), or of the *tangent* to a *curve*, or the area contained by a curve. The reader may find it helpful to translate this geometrical language into rough sketches, for a graph may serve in analysis as illustration, provided illustration is not mistaken for demonstration. But a diagram in the text catches the eye and tends to bring into play an oft-misleading geometrical intuition, before there has been time to follow the argument, with a con-

sequent failure to see the need for a detailed analysis of what is deemed self-evident in the picture.

The text of the book covers the Calculus Syllabus for the General Honours degree of London and the provincial universities, and the London special degree. The first eleven chapters contain all that is needed of the Calculus for the ordinary and advanced papers in the Higher School Certificate examination. The order of chapters may be varied slightly to suit individual requirements. For instance Chapter XI, on differential equations, may be taken before Chapter X, which deals with area and arc-length and curvature in a plane. A number of theorems have been marked with an asterisk, and these may be omitted at a first reading.

In addition to the 'drill' examples there are over four hundred harder examples, with complete solutions. The solved examples are of various types. Some play an integral part in the development of the book, and should be regarded as additions to the chapter whose number they bear. Others indicate alternative lines of development, or generalizations and extensions, of the theorems in the text; the majority of the examples, however, are exercises in the technique of the Calculus.

In the preparation of this book most of the standard works on Analysis were consulted on one point or another, and although almost every demonstration has some element of novelty in it, the debt owed to these works is a great one. To Professor E. H. Neville my warmest thanks are due for many valuable suggestions, and for his generous and unstinted help in correcting the proofs. I should like to express also my gratitude to the Delegates and Staff of the Clarendon Press, and my pleasure in the excellence of their work.

I am indebted, above all, to Professor T. A. A. Broadbent, who read the manuscript and criticized the text down to the smallest detail. From the choice of words to the order of development, every aspect of the book has benefited from his penetrating observations and helpful suggestions, and I welcome this opportunity to thank him and to acknowledge how much I owe both to his kindness and to his judgement, knowledge, and experience.

The 'Uniform Calculus' is not an attempt to solve philosophical problems in the foundations of mathematics. The selective modifications which have been made are made solely with a view to

simplifying the technique, not to meet the well-known philosophical objections to the current system. Though no use is made of proof by *reductio ad absurdum*, the Uniform Calculus is nevertheless not strictly finitist, for many of the arguments are based upon undecidable disjunctions; for instance the proof of the fundamental theorem on the determination of the limit of a convergent sequence, and the proof that a continuous function which takes both positive and negative values takes also the value zero (though the proof is finitist if the function is rational). As far as possible, however, the emphasis has been placed on *practicable constructions*, and in this respect at least the present approach may be philosophically preferable to the Dedekind-Russell formulation.

R. L. G.

THE UNIVERSITY,
READING

CONTENTS

I

SEQUENCE AND SERIES

THE DECIMAL NOTATION. LIMIT OF A CONVERGENT SEQUENCE. PROPERTIES OF POWER SERIES

In elementary arithmetic we meet several kinds of numbers, the natural numbers we use for counting, the fractions which record a selection from a total, like 5 out of 8, and the positive and negative integers which record increase and decrease. The natural numbers alone, however, are fundamental, for all calculations with fractions or integers are based upon calculations with natural numbers. For example, to find the product of two fractions we multiply the natural numbers which are the numerators and denominators of the fractions, and to find the sum of two positive integers, like $+2$ and $+3$, we add the natural numbers 2 and 3. The familiar numerals 2 to 9 are, of course, only abbreviations for the natural numbers 'one and one', 'one and one and one', and so on, and such numerals as 10, 25, or 173 are also abbreviations, but in a different way, for these are numerals in a scale notation in which the digits represent so many units, tens, hundreds, etc., according to their position from right to left.

The expression of numbers in a scale serves to facilitate calculation; to add or multiply in the scale of ten, for instance, we need but tabulate the sums and products of any two of the digits one to nine, and in the scale of two only the two results $1+1 = 10$, $1\times 1 = 1$ are needed. The smaller the scale the smaller the addition and multiplication tables, but this advantage of a small scale is offset by its failure to represent large numbers by means of comparatively few digits; for example, in the scale of two the number ninety-nine is a number of seven digits whereas in the scale of ten it has but two.

The convenience of a scale notation was extended to fractions by the introduction of the decimal fraction, which marks one of the great steps in the development of mathematics. A decimal fraction (or, shortly, a decimal) records a selection from a total by means of single number, like ·702, the dot serving to distinguish ·702 from 702 and the number of digits alone showing that the

selection is made from a total of 10^3 (the scale number 10 being taken for granted).

The decimal representation of fractions is, however, not quite as simple as this suggests. A fraction like $\frac{1}{3}$ is not represented by any terminating decimal, since there is a remainder unity when 3 is divided into any power of 10; to express the fact that the remainder recurs we write a dot over the partial quotient, so that $\frac{1}{3}$ is represented by $\cdot\dot{3}$. $\cdot\dot{3}$ does not stand for any of the terminating decimals $\cdot 3$, $\cdot 33$, $\cdot 333$, etc., but is a new number sign. We might say that the sign $\cdot\dot{3}$ stands for the *endless* decimal each of whose digits is a three.

Although we first meet examples of endless decimals in attempting to find a decimal representation for certain fractions, it has long been known that there are endless decimals which do not represent any fraction so that, in general, an endless decimal is a new kind of number; for example, the endless decimal which arises from the familiar process of extracting the square root of 2 is not the decimal representation of any fraction, for (in a sense of which we shall have more to say shortly) the square of this endless decimal has the value 2 whereas it can quite easily be shown that the square of any fraction p/q differs from 2 by at least $1/q^2$.

Since an endless decimal consists of an *endless* succession of digits we are obliged to identify it by a description or a name; we have already met instances of this, such as the specification of the endless decimal, each digit of which is a 3, by the sign $\cdot\dot{3}$, and another familiar example is afforded by the endless decimal which expresses the number of radians in an angle of 180° and which is known as 'π'. In addition to specific endless decimals, named and identified, we can consider the *general idea* of a decimal as an endless succession of digits.

The step from the terminating to the endless decimal is the step from Arithmetic to the Calculus. The arithmetical operations with endless decimals are of an essentially different character from the operations with natural numbers or terminating decimals. In elementary arithmetic, if we wished to multiply the endless decimals $\cdot\dot{3}$ and $\cdot\dot{6}$ we should multiply the fractions $\frac{1}{3}$ and $\frac{2}{3}$ which these endless decimals represent and should call $\cdot\dot{2}$, which represents $\frac{2}{9}$, the product of the decimals. Since, however, as we have

observed, there are endless decimals which do not represent fractions, this method of multiplying endless decimals is not generally applicable. If we consider in turn the products $\cdot 3\times\cdot 6 = \cdot 18$, $\cdot 33\times\cdot 66 = \cdot 2178$, $\cdot 333\times\cdot 666 = \cdot 221778$, $\cdot 3333\times\cdot 6666 = \cdot 22217778$ we perceive that, step by step, we are constructing the decimal $\cdot 2222\ldots$. We should have little difficulty in convincing ourselves that the more 3's and 6's in the multiplier and multiplicand the longer will be the run of 2's in the product, i.e. the more digits, of the product will become fixed. From the second product onwards the first digit is a 2, from the third product onwards the second digit also attains its final value, and so on. But if we seek to *prove* that this is so, we cannot look for the proof in the products themselves, for however many multiplications we carry out there will always remain multipliers and multiplicands with more digits and so, for aught we have shown, we might reach products which do not commence with a run of 2's. That the successive digits of the products do in fact gradually settle down is proved by showing that *every* product lies between certain values. Since each of the numbers $\cdot 3$, $\cdot 33$, $\cdot 333,\ldots$ is greater than its predecessor, and so, too, each of the numbers $\cdot 6$, $\cdot 66$, $\cdot 666,\ldots$, therefore each of the products $\cdot 3\times\cdot 3$, $\cdot 33\times\cdot 66$, $\cdot 333\times\cdot 666,\ldots$ is greater than its predecessor; moreover, if we choose any of the numbers, say $\cdot 333$, and add 1 to its final digit, we obtain a number, $\cdot 334$ in fact, which is greater than any of the decimals whose digits are all 3's, and similarly $\cdot 667$ is greater than any decimal whose digits are all 6's, so that each of the products $\cdot 3\times\cdot 6$, $\cdot 33\times\cdot 66$, $\cdot 333\times\cdot 666$, $\cdot 3333\times\cdot 6666,\ldots$ is less than $\cdot 334\times\cdot 667$. Thus each of the products, from the third onwards, lies between $\cdot 333\times\cdot 666 = \cdot 221778$ and $\cdot 334\times\cdot 667 = \cdot 222778$ and so each product necessarily commences with the digits $\cdot 22$. In this way we can determine properties of a succession of products without carrying out the actual multiplications.

If we denote the numbers $\cdot 3$, $\cdot 33$, $\cdot 333,\ldots$ by $a_1, a_2, a_3,\ldots$ and the numbers $\cdot 6$, $\cdot 66$, $\cdot 666,\ldots$ by $b_1, b_2, b_3,\ldots$ we may exhibit the proof that the digits of the products $a_1 b_1$, $a_2 b_2$, $a_3 b_3,\ldots$ settle down, in a more general light. Since each decimal a_{n+1} is formed by adding a digit to the preceding a_n, it follows that for any p the number $a_{n+p}-a_n$ commences with n zeros and accordingly

$$a_{n+p}-a_n < 1/10^n;$$

for example $\cdot 333333 - \cdot 33 = \cdot 003333 < \cdot 01.$

Similarly $$b_{n+p} - b_n < 1/10^n$$

and therefore

$$a_{n+p}b_{n+p} < (a_n + 1/10^n)(b_n + 1/10^n)$$
$$= a_n b_n + (a_n + b_n)/10^n + 1/10^{2n}.$$

Furthermore, for any n, $a_n < \cdot 4$ and $b_n < \cdot 7$ and so

$$a_{n+p}b_{n+p} - a_n b_n < (11/10 + 1/10^n)/10^n < 2/10^n < 1/10^{n-1}.$$

This shows that (unless the $(n-1)$th digit of the product $a_n b_n$ is a 9) the product $a_{n+p}b_{n+p}$ has the same first $n-2$ digits as $a_n b_n$. E.g. if $n = 4$, since $a_4 b_4 = \cdot 22217778$, therefore $a_{p+4}b_{p+4}$, for any p, lies between $\cdot 22217778$ and $\cdot 22317778$, i.e. all the products $a_n b_n$, after the fourth, commence with $\cdot 22$.

The property of the sequences $a_1, a_2, a_3, \ldots, b_1, b_2, b_3, \ldots$, and $a_1 b_1, a_2 b_2, a_3 b_3, \ldots$, which can be deduced from the inequalities $a_{n+p} - a_n < 1/10^n$, $b_{n+p} - b_n < 1/10^n$, and $a_{n+p}b_{n+p} - a_n b_n < 1/10^{n-1}$, that in each sequence the difference between the nth term and any succeeding term can be made as small as we please by choosing n sufficiently great, is called *convergence*. Convergence is a sequence-property of fundamental importance in the calculus and we shall now define it formally.

1. Definition of convergence

A sequence of terminating decimals $u_1, u_2, u_3, \ldots$ is said to be convergent if we can find a sequence $n_1, n_2, n_3, \ldots$ such that for any r and $N \geqslant n_r$, the positive value of the difference between u_N and u_{n_r} is less than $1/10^r$.

1.01. In less formal language: The sequence $u_1, u_2, u_3, \ldots$ is convergent if we can make the difference between the term u_n and *any* term which follows it less than any chosen number by a suitable choice of the value of n. In general the requisite value of n will depend upon the desired difference between u_n and its successors, though of course if from some term onwards the terms of the sequence are all equal the convergence condition is satisfied for a single value of n, no matter how small the required difference between u_n and its successors may be.

EXAMPLE. The sequence whose nth term is $(\surd n-1)/\surd n$ is convergent, because, for $N > n$,

$$(\surd N-1)/\surd N-(\surd n-1)/\surd n = 1/\surd n-1/\surd N < 1/\surd n,$$

and so, if $n_r = 10^{2r}$, then

$$(\surd N-1)/\surd N-(\surd n_r-1)/\surd n_r < 1/10^r$$

for any N not less than n_r.

1.02. The decimal notation admits a certain ambiguity, since the two expressions $\cdot\dot{9}$, $1\cdot\dot{0}$ represent the same number; that this is the case follows from the fact that if we divide unity into $1\cdot\dot{0}$ we may write the result either as $1\cdot\dot{0}$ or as $\cdot 9$ with remainder 1, which gives rise to the quotient $\cdot\dot{9}$. Accordingly, if $a_m < 9$, the expressions $a_0\cdot a_1 a_2 \ldots a_m \dot{9}$ and $a_0\cdot a_1 a_2 \ldots a_{m-1} a_m^*$, $a_m^* = a_m+1$, represent the same number, for if we multiply the first by 10^m we obtain $10^m a_0 + a_1 a_2 \ldots a_m \cdot \dot{9}$ which equals $10^m a_0 + a_1 a_2 \ldots a_{m-1} a_m^*$, which is 10^m times the second.

1.1. The limit of a convergent sequence of terminating decimals

All numbers which differ from some terminating decimal a by less than $1/10^{n+1}$ have, in general, the same whole part and first n decimal figures as a itself; for instance all the numbers which differ from $13\cdot 27825$ by less than $1/10^4$ lie between $13\cdot 2781$ and $13\cdot 2783$, and so all commence with the digits $13\cdot 278$, and $13\cdot 278$ has the same whole part and first three digits as $13\cdot 27825$.

Thus if s_n is a convergent sequence, so that for a certain n_r, and $n \geqslant n_r$, all s_n differ from s_{n_r} by less than $1/10^r$, then all s_n $(n \geqslant n_r)$ have the same whole part and first $r-1$ decimal figures as s_{n_r} itself. Let l_0 be the whole part of s_{n_1}, so that l_0 is also the whole part of s_n for all $n \geqslant n_1$, and in particular the whole part of s_{n_2} (since $n_2 \geqslant n_1$). Hence if l_1 is the first decimal figure of s_{n_2}, then $l_0\cdot l_1$ has the same whole part and first decimal figure as s_{n_2}, and therefore the same whole part and first decimal figure as all s_n after s_{n_2}, and in particular the same as s_{n_3}. Wherefore if l_2 is the second decimal figure of s_{n_3}, then $l_0\cdot l_1 l_2$ has the same whole part and first two decimal figures as s_{n_3} and therefore the same as s_n for all n after n_3, and so on. Accordingly if l_p is the pth decimal figure of $s_{n_{p+1}}$, then the *endless* decimal $l_0\cdot l_1 l_2 \ldots l_p \ldots$

has the same whole part and first p decimal figures as all s_n from $s_{n_{p+1}}$ onwards.

The decimal $l_0{\cdot}l_1 l_2...$ constructed in this way is called the LIMIT of the convergent sequence s_n, and s_n is said to *tend* to l. To determine the limit to p decimal places we simply delete all the figures in $s_{n_{p+1}}$ after the pth decimal place.

By construction, the limit of a convergent sequence s_n is an endless decimal which has as many digits as we please in common with all s_n from some suitable value of n onwards.

If l is the limit of s_n we write for short '$\lim s_n = l$' or '$s_n \to l$'. The expression '$\lim s_n = l$' is read 'the limit of s_n is l', and '$s_n \to l$' as 's_n tends to l'.

A difficulty may arise if the $(p+1)$th digit of $s_{n_{p+1}}$ is a zero or a nine, since in this case the pth digit of an s_n after $s_{n_{p+1}}$ may not be the same as the pth digit of $s_{n_{p+1}}$. For instance $\cdot 370004$ and $\cdot 369997$ differ by less than $1/10^5$ but have different figures in the second, third, and fourth decimal places.

It follows that it may happen, exceptionally, that when we choose the whole part and first p decimal figures of $s_{n_{p+1}}$, taking $p = 0, 1, 2,...$ in turn, instead of obtaining a succession of the form $l_0, l_0{\cdot}l_1, l_0{\cdot}l_1 l_2, l_0{\cdot}l_1 l_2 l_3,...$, the numbers we obtain may alternate like

$$2, \quad 1{\cdot}9, \quad 2{\cdot}00, \quad 1{\cdot}999, \quad 2{\cdot}000, \quad ....$$

Here the digits do not follow on, but none the less the successive terms have more and more figures in common with the number 2 expressed in one or other of the possible forms $2{\cdot}\dot{0}$ or $1{\cdot}\dot{9}$. In such a case we call 2 the limit and observe that we can still say the limit has as many digits in common with s_n as we please provided we admit the dual representation of the limit as $2{\cdot}\dot{0}$ and as $1{\cdot}\dot{9}$.

EXAMPLES. 1. The sequence $\cdot 5, \cdot 42, \cdot 415, \cdot 4142, \cdot 41415, \cdot 414142,...$ in which the $(2n)$th term is the decimal (with $2n$ digits) $\cdot 4141...4142$, and the $(2n+1)$th term is the decimal (with $2n+1$ digits) $\cdot 4141...415$, is convergent, with limit $\cdot\dot{4}\dot{1} = \cdot 414141...$, which has the same first digit as all the terms after the second, the same first and second digits as all the terms after the third, and so on.

2. The sequence $\cdot 3, \cdot 191, \cdot 21, \cdot 1991, \cdot 201, \cdot 19991, \cdot 2001, \cdot 199991, \cdot 20001,...$ in which the $(2n)$th term is the decimal (with n nines) $\cdot 199...991$, and the $(2n+1)$th term is the decimal (with $n-1$ zeros)

$\cdot 200\ldots001$, is convergent, with limit $\cdot 2 = \cdot 1\dot{9}$; the even terms, from the $(2n)$th onwards, have the same first $n+1$ digits as the limit, in the representation $\cdot 1\dot{9}$, and from the $(2n+1)$th onwards, the odd terms have the same first n digits as the limit, in the representation $\cdot 2\dot{0}$.

3. Prove that the sequence

$$s_1 = 1, \quad s_2 = s_1+1, \quad s_3 = s_2+1/2^2,$$
$$s_4 = s_3+1/3^3, \quad \ldots, \quad s_{n+1} = s_n+1/n^n, \quad \ldots$$

is convergent, and evaluate the limit to three places of decimals.

We prove first that

$$0 < s_{r+n}-s_r \leqslant \frac{1}{r^r}\left(2-\frac{1}{r^{n-1}}\right).$$

These inequalities hold with $n = 1$, by definition; if they are true for $n = k$, then since

$$s_{r+k+1}-s_r = (s_{r+k+1}-s_{r+k})+(s_{r+k}-s_r)$$

we have

$$0 < s_{r+k+1}-s_r \leqslant \frac{1}{(r+k)^{r+k}}+\frac{1}{r^r}\left(2-\frac{1}{r^{k-1}}\right)$$
$$< \frac{1}{r^{r+k}}+\frac{1}{r^r}\left(2-\frac{1}{r^{k-1}}\right) = \frac{2}{r^r}-\frac{r-1}{r^{r+k}}$$
$$\leqslant \frac{2}{r^r}-\frac{1}{r^{r+k}} \quad (r \geqslant 2),$$

so that the inequalities hold also for $n = k+1$, wherefore, by induction, they are true for all values of n.

It follows that

$$0 < s_{r+n}-s_r < 2/r^r \leqslant 1/r^{r-1} \quad (r \geqslant 2),$$

and so, changing r into $r+1$,

$$0 < s_N-s_{r+1} < 1/(r+1)^r < 1/10^r \quad (N \geqslant r+1 \geqslant 10),$$

which proves that the sequence s_n is convergent.

We readily calculate

$$s_2 = 2, \quad s_3 = 2\cdot 25, \quad s_4 = 2\cdot 287037\ldots,$$
$$s_5 = 2\cdot 290943\ldots, \quad s_6 = 2\cdot 291263\ldots.$$

Since $$0 < s_n-s_6 < 2/6^6 < 1/10^4 \quad (n > 6),$$

therefore all s_n after s_6 lie between $2\cdot 291263\ldots$ and $2\cdot 291363\ldots$ so that $s_n = 2\cdot 291$ to three places of decimals, for all $n \geqslant 6$, whence $\lim s_n = 2\cdot 291$ to three places of decimals.

1.2. In simple cases it is possible to estimate the limit of a sequence, by inspection, before proving that the sequence is convergent. Suppose that for some sequence $s_1, s_2, s_3, \ldots$ *we find a number l, and a sequence n_r*, such that

$$|l-s_n| < 1/10^r$$

for all n not less than n_r, then the sequence s_n is convergent and $\lim s_n = l$. For

$$|s_n - s_{n_r}| = |(s_n - l)+(l-s_{n_r})| \leqslant |l-s_n|+|l-s_{n_r}| < 2/10^r < 1/10^{r-1} \quad (n \geqslant n_r),$$

which proves that s_n is convergent.

If l is a terminating decimal,

$$a_0 \cdot a_1 a_2 \ldots a_p = a_0 \cdot a_1 a_2 \ldots a_{p-1} b_p \dot{9}$$
$$(a_p > 0,\ b_p = a_p - 1),$$

then, when $n \geqslant n_{p+q}$, since $|l-s_n| < 1/10^{p+q}$, all s_n lie between $a_0 \cdot a_1 a_2 \ldots a_{p-1} b_p 99\ldots9$, with q nines, and $a_0 \cdot a_1 a_2 \ldots a_p 00\ldots01$, with $q-1$ zeros, and therefore s_n has the same whole part and first $p+q-1$ decimal figures as l (in one or other decimal representation); thus l is the limit of s_n. And if l is a non-terminating decimal $a_0 \cdot a_1 a_2 \ldots$, then we can determine M as great as we please such that a_M is not zero. If $a_M = 9$ we can find $N > M$ such that $a_N < 9$ (since l does not terminate with a recurrence of nines), and if $a_M < 9$ we take $N = M$; but $|l-s_n| < 1/10^N$ $(n \geqslant n_N)$, and so all s_n $(n \geqslant n_N)$ have the same whole part and first $M-1$ digits as l.

This shows that l is the limit of the sequence.

Examples. 1. Since

$$|(n+1)/n - 1| = 1/n < 1/10^r \quad (n > 10^r),$$

it follows that 1 is the limit of the sequence

$$2/1,\ 3/2,\ 4/3,\ \ldots,\ (n+1)/n,\ \ldots$$

and the sequence is convergent.

2. The sequence $1, k, k^2, \ldots, k^n, \ldots$ is convergent, with limit zero, if $-1 < k < 1$, and convergent with limit unity if $k = 1$; if $k > 1$ or $k \leqslant -1$ the sequence is not convergent.

We distinguish the following cases:

(i) $k = 0$. Here $k^n = 0$ for all n and so $\lim k^n = 0$.

$k = 1$. $k^n = 1$ for all n and therefore $\lim k^n = 1$.

(ii) $0 < k < 1$. Write $l = 1/k$, so that $l > 1$ and let $a = l-1 > 0$. Then $l^n = (1+a)^n > na$ [since $(1+a)^n$ is the sum of positive terms $1+na+\binom{n}{2}a^2+...$; or alternatively by induction, proving $(1+a)^n > 1+na$ by means of the inequalities

$$(1+a)(1+na) = 1+na+a+a^2 > 1+(n+1)a],$$

and so if $n \geqslant 10^r/a$ it follows that $l^n > 10^r$. Hence $k^n < 1/10^r$ provided that $n \geqslant 10^r/a$, which proves that $\lim k^n = 0$.

(iii) $-1 < k < 0$. Then $0 < |k| < 1$ and so $|k|^n < 1/10^r$ by (ii), whence $|k^n| < 1/10^r$ and so $\lim k^n = 0$.

(iv) $k > 1$. Let $a = k-1$, so that $k^p = (1+a)^p > pa > 2$, provided $p > 2/a$. Thus $k^{n+p}-k^n = k^n(k^p-1) > 1$, for any n whatsoever, provided $p > 2/a$, and therefore the sequence k^n is not convergent.

(v) $k < -1$. Here $|k| > 1$ and so

$$|k^{n+2p}-k^n| = |k|^n|k^{2p}-1| = |k|^n||k|^{2p}-1| > 1,$$

provided $p > 1/a$, and therefore the sequence k^n is not convergent.

(vi) $k = -1$. Since $|k^{n+1}-k^n| = |k|^n|k-1| = 2$, therefore k^n is not convergent.

It may assist the reader to follow the proofs we have given that certain sequences are *not* convergent if we formulate the negation of the convergence property in detail. A sequence s_n is *not* convergent if we can find a (fixed) k and a sequence $p_r > r$ such that, for any r,

$$|s_{p_r}-s_r| \geqslant 1/10^k.$$

In less formal language: s_n is not convergent if *any* s_r is followed by *some* s_n which differs from it by more than a certain positive quantity. The value of k in the inequality $|s_{p_r}-s_r| \geqslant 1/10^k$ is constant and the inequality must hold for this fixed k whatever value r may take. Furthermore there may be more than one value of p_r (for each r) for which the inequality holds, but it is sufficient if there is a *single* value of p_r (for each r) for which the inequality is known to be satisfied.

Consider in turn the cases (iv), (v), and (vi) above. In (iv) we showed that, for any n, k^n differed from *any* following k^{n+p} by more than unity, whereas in (v) what we proved was that every

k^n differed from *certain* following terms, namely the terms k^{n+2p}, by more than unity. In (vi) we were satisfied to prove that each k^n differed from the *single* term which follows it by as much as 2.

To prove convergence we must show that the difference between s_{n_r} and *every* s_n after it is less than $1/10^r$; lack of convergence requires only that each s_n differs from *some one* subsequent term by more than fixed amount. The negative of convergence is generally called *divergence*.

1.3. Notation

We shall have such frequent occasion to consider numbers whose positive value is less than some power of 1/10, that we shall denote such numbers by a special abbreviation. We shall write '$0(r)$' for any number which is zero to r places of decimals, i.e. for any number whose positive value is less than $1/10^r$. The equation '$x = 0(r)$', which may be read as 'x is 0 to r places', is the equivalent of '$|x| < 1/10^r$'.

Since the sum or difference of any two numbers which are zero to at least r places is zero to at least $r-1$ places, we write

$$0(r)+0(s) = 0(r-1), \quad \text{where } s \geqslant r,$$

and since the product of a number less than $1/10^r$ with a number less than $1/10^s$ is less than $1/10^{r+s}$ we write $0(r).0(s) = 0(r+s)$. If $r_1, r_2,..., r_n$ are all not less than r, and if $n \leqslant 10$, then

$$1/10^{r_1}+1/10^{r_2}+...+1/10^{r_n} \leqslant n/10^r \leqslant 10/10^r = 1/10^{r-1},$$

and therefore, provided $n \leqslant 10$,

$$0(r_1)+0(r_2)+...+0(r_n) = 0(r-1).$$

In this notation the condition that a sequence $u_1, u_2, u_3,...$ is convergent is expressed by:

$$\text{for } N \geqslant n_r, \quad u_N-u_{n_r} = 0(r) \quad \text{or} \quad u_N = u_{n_r}+0(r),$$

and the condition that l is the limit of the sequence is:

$$\text{for } n \geqslant n_r, \quad u_n = l+0(r).$$

Observe that if $a-b = 0(r)$ then $b-a = 0(r)$, since each equation says that $|a-b| < 1/10^r$.

Whether x is a terminating or endless decimal we denote the decimal formed by the whole part and first r decimal digits of x by $(x)_r$. Thus $x = 0(r)$ and $(x)_r = 0$ say the same thing. For any decimal x, $x-(x)_r$ commences with r zeros and so $x-(x)_r = 0(r)$.

The condition that a sequence a_n is convergent takes the simple form:

$$\text{for all } r, \quad (a_n - a_{n_r})_r = 0 \quad (n \geqslant n_r);$$

and the relation between a convergent sequence a_n and its limit l is

$$(l)_r = (a_n)_r \quad \text{for all } r \text{ and } n \geqslant n_r.$$

Note that if $p \geqslant r$ then $((x)_p)_r = (x)_r$, and if x is a terminating decimal of less than $r+1$ decimal digits, then $(x)_r = x$.

1.31. In order to prove, for some U_n, that

$$U_n = 0(r) \quad \text{for all } n \geqslant n(r)$$

it is sufficient to prove

$$U_n = A0(r) \quad \text{for all } n \geqslant \nu(r)$$

provided that, when the expression A depends upon n or r, there is a constant K such that $|A| \leqslant K$ for any n and r, i.e. provided that A is *bounded*. For if p is chosen so that $10^p \geqslant K$ and if $n \geqslant \nu(p+r)$, then

$$|U_n| < |A|/10^{p+r} \leqslant 1/10^r,$$

and so, taking $n(r) = \nu(p+r)$, we have

$$U_n = 0(r) \quad \text{for } n \geqslant n(r).$$

1.4. Sums and products of convergent sequences of terminating decimals

1.41. If the sequences a_n and b_n are convergent then the sequence a_n+b_n is convergent.

For we can find m_r and n_r so that for $M \geqslant m_r$ and $N \geqslant n_r$,

$$a_M = a_{m_r}+0(r), \qquad b_N = b_{n_r}+0(r).$$

Let λ_r denote the greater of the two numbers m_r, n_r, and therefore, if $k \geqslant \lambda_r$,

$$\begin{aligned}(a_k+b_k)-(a_{\lambda_r}+b_{\lambda_r}) &= (a_k-a_{\lambda_r})+(b_k-b_{\lambda_r}) \\ &= (a_k-a_{m_r})+(a_{m_r}-a_{\lambda_r})+(b_k-b_{n_r})+(b_{n_r}-b_{\lambda_r}) \\ &= 0(r)+0(r)+0(r)+0(r) \\ &= 0(r-1).\end{aligned}$$

Thus the sequence a_n+b_n is convergent.

1.411. Taking $-b_n$ for b_n we see also that a_n-b_n is convergent.

1.42. We show next that the sequence $a_n b_n$ is convergent.

Choosing λ_r as above we find

$$a_k b_k = \{a_{\lambda_r}+0(r-1)\}\{b_{\lambda_r}+0(r-1)\}$$
$$= a_{\lambda_r} b_{\lambda_r}+\{a_{\lambda_r}+b_{\lambda_r}+0(r-1)\}0(r-1).$$

Now for any $m \geqslant m_1$ and any $n \geqslant n_1$

$$a_m = a_{m_1}+0(1), \qquad b_n = a_{n_1}+0(1)$$

and so if we choose p so that $|a_{m_1}|$ and $|a_{n_1}|$ are both less than p, it follows that for any m greater than both m_1 and n_1 we have

$$|a_m| < p+1 \quad \text{and} \quad |b_m| < p+1$$

and therefore $\qquad |a_m+b_m+0(r-1)| < 2p+3,$

and so the sequence $a_n b_n$ is convergent (by 1.31).

1.43. We can also show that $1/a_n$ is convergent, provided that for some fixed $q > 0$, $|a_n| > q$ for all n.

For $\quad |1/a_N-1/a_n| = |a_n-a_N|/|a_n||a_N| < \dfrac{1}{q^2}|a_n-a_N|$

and $|a_N-a_n| = 0(r)$ for $N \geqslant n$ and a suitable value of n.

1.431. From 1.42 and 1.43 we deduce the convergence of the sequence b_n/a_n, provided $|a_n| > q$ for some fixed $q > 0$, and all n.

1.44. Arithmetical operations with *endless* decimals

Definition of sum and product of endless decimals. If a_n and b_n are the values to n decimal places of two endless decimals α and β, then since a_N-a_n and b_N-b_n are both less than $1/10^n$, a_n and b_n are convergent and therefore, by 1.41, a_n+b_n is convergent. Let s be the limit of the sequence a_n+b_n; then we define s to be the *sum* of the endless decimals α and β. In the notation of § 1.3, $\alpha+\beta = \lim\{(\alpha)_n+(\beta)_n\}$. If d is the limit of the convergent sequence a_n-b_n we define d to be the *difference* of α and β, i.e.

$$\alpha-\beta = \lim\{(\alpha)_n-(\beta)_n\}.$$

Furthermore, if p is the limit of the convergent sequence $a_n b_n$ and q the limit of a_n/b_n (supposing that $|b_n| \geqslant \lambda > 0$ for all n), then p and q are defined to be product and quotient respectively of the endless decimals α and β, i.e.

$$\alpha\beta = \lim(\alpha)_n.(\beta)_n \quad \text{and} \quad \alpha/\beta = \lim(\alpha)_n/(\beta)_n.$$

1.441. A sequence of *endless* decimals $s_1, s_2, s_3,\ldots$ is said to be convergent if $s_n-s_{n_r} = 0(r)$, $n \geqslant n_r$.

This definition is the same as for a sequence of terminating decimals; observe, however, the convergence condition involves the *difference* in terms of the sequence, and the difference of two *endless* decimals itself involves the convergence of a sequence of terminating decimals.

1.4411. If s_n is a convergent sequence of endless decimals then the sequence $(s_n)_n$ converges.

For

$$\begin{aligned}(s_N)_N-(s_n)_n &= \{s_N-s_n\}+\{s_n-(s_n)_n\}-\{s_N-(s_N)_N\}\\ &= 0(r)+0(n)+0(N), \quad \text{if } N \geqslant n \geqslant n_r,\\ &= 0(r-1), \quad \text{if } n \geqslant r.\end{aligned}$$

1.4412. If s_n is convergent then s_n converges to a limit l; for if l is the limit of the convergent sequence of *terminating* decimals $(s_n)_n$ then $(l)_k = \{(s_n)_n\}_k$, $n \geqslant n_k$, and so if $n > k$, so that

$$\{(s_n)_n\}_k = (s_n)_k,$$

we have $(l)_k = (s_n)_k$, so that l is the limit of the sequence s_n.

1.442. Inequalities. If x and y are endless decimals (not terminating in a recurrence of 9's) and if for some n, $(x)_n > (y)_n$, we write $x > y$; if $(x)_n < (y)_n$, we write $x < y$, and if $(x)_n = (y)_n$ for *all* values of n, then $x = y$.

We observe that if $(x)_r = 0$, for some r, then $|x| < 1/10^r$, for if $(x)_r = 0$, then $(-1/10^r)_r < (x)_r < (1/10^r)_r$, and so

$$-1/10^r < x < 1/10^r.$$

1.45. The limit of the sum of two convergent sequences is the sum of the limits of the sequences.

Suppose $\lim a_n = \alpha$, $\lim b_n = \beta$, and $\lim(a_n+b_n) = s$, so that for $n \geqslant$ some k, $\alpha = a_n+0(r)$, $\beta = b_n+0(r)$, and $s = a_n+b_n+0(r)$, and therefore $\alpha+\beta = s+0(r-1)$. This equation is true for any value of r and therefore $\alpha+\beta$ and s are equal to r places for any r, whence it follows that $\alpha+\beta = s$.

1.451. Replacing b_n by $-b_n$ and β by $-\beta$ it follows from 1.45 that the limit of the difference of two convergent sequences is the difference of the limits of the sequence.

1.452. The limit of the product of two convergent sequences is the product of the limits of the sequences.

For $\alpha = a_n+0(r)$, $\beta = b_n+0(r)$, $p = a_n b_n+0(r)$ and so

$$\alpha\beta-p \doteq (a_n+b_n).0(r)+0(2r).$$

But $a_n+b_n = \alpha+\beta+0(r-1)$ and so $\alpha\beta-p = (\alpha+\beta)0(r)+0(2r-2)$. Choosing m so that $(\alpha+\beta) < 10^m$ it follows that

$$|\alpha\beta-p| < 1/10^{r-m-1},$$

for any r, and so $\alpha\beta = p$.

1.453. If a sequence a_n converges to a non-zero limit α then $1/a_n$ converges to $1/\alpha$. For $a_n = \alpha+0(k)$ and so, choosing k so that $\frac{1}{2}|\alpha| > 1/10^k$, it follows that $|a_n| > |\alpha|-\frac{1}{2}|\alpha| > 1/10^k$.

Hence

$$|1/\alpha-1/a_n| = |a_n-\alpha|/|\alpha||a_n| < |a_n-\alpha|10^{2k}$$

and so if n is chosen great enough to ensure $a_n = \alpha+0(r+2k)$,

$$|1/\alpha-1/a_n| < 1/10^r,$$

which shows that $\lim 1/a_n = 1/\alpha$.

Example. To find the limit of the sequence $(n+1)/(2n+3)$.

Consider $1+1/n$. Since $1/n = 0(r)$ for $n \geqslant 10^{r+1}$, $1/n \to 0$, and therefore (as the limit of the sequence 1, 1, 1, 1,... is 1) $1+1/n \to 1$. Similarly $2+3/n \to 2$ and so $(1+1/n)/(2+3/n) \to 1/2$. But

$$(1+1/n)/(2+3/n) = \left(\frac{n+1}{n}\right)\Big/\left(\frac{2n+3}{n}\right) = (n+1)/(2n+3).$$

This example illustrates an important technique. Neither of the sequences $n+1$, $2n+3$ is convergent, since for instance

$$N+1-(n+1) = N-n > 1, \quad \text{for } N > n,$$

however great n may be, but by dividing the numerator and denominator of the fraction $(n+1)/(2n+3)$ by n, the given sequence is seen to be the quotient of the two convergent sequences

$$1+1/n, \qquad 2+3/n.$$

The sequence $(n^2+2n+2)/(n+1)$ is not convergent, for

$$(n^2+2n+2)/(n+1) = (n+1)+1/(n+1),$$

and $n+1$ is not convergent whereas $1/(n+1)$ converges. (If $(n+1)+1/(n+1)$ were convergent then the difference between it and the convergent sequence $1/(n+1)$ would also be convergent, but this difference is $n+1$.) Dividing the numerator and denominator of $(n^2+2n+2)/(n+1)$ by n^2 we obtain

$$(1+2/n+2/n^2)/(1/n+1/n^2);$$

since $1/n \to 0$, therefore the product $1/n \times 1/n \to 0$ and so

$$1+2/n+2/n^2 \to 1, \quad \text{and} \quad 1/n+1/n^2 \to 0.$$

This does not however imply the convergence of

$$(n^2+2n+2)/(n+1),$$

for Theorem 1.453 requires that the limit of the denominator be other than zero.

1.46. If a_n is a convergent sequence such that $\lim a_n = \alpha$ and if $a_n \geqslant \lambda$, for any n, then $\alpha \geqslant \lambda$.

Since $a_n \to \alpha$, $\alpha > a_n - 1/10^r$ for a sufficiently great n, and so $\alpha > \lambda - 1/10^r$ for any r. But by choosing r great enough we can make $\lambda - 1/10^r$ greater than *any* number less than λ and so α is greater than every number less than λ. Since α is greater than any number less than λ it is not itself one of the numbers less than λ and so α is greater than or equal to λ. Alternatively, for an appropriate n, $(\alpha)_r = (a_n)_r \geqslant (\lambda)_r$ and so $\alpha \geqslant \lambda$.

1.47. If a_n is a convergent sequence such that $\lim a_n = \alpha$, and if $a_{n+1} > a_n$ for any n, then $\alpha > a_n$ for any n.

Since $a_{n+1} - a_n$ is positive we can choose r, depending on n, such that $1/10^r < a_{n+1} - a_n$, and since $a_n \to \alpha$, $a_m < \alpha + 1/10^r$ for any $m \geqslant$ some m_r. Therefore for any n, $a_m < \alpha + (a_{n+1} - a_n)$ provided m is great enough, and so $a_n + (a_m - a_{n+1}) < \alpha$; but when $m > n+1$, then $a_m > a_{n+1}$ and therefore $a_n < \alpha$, and this is true for any n.

Similarly, if $a_n \to \alpha$ and $a_{n+1} < a_n$ then $a_n > \alpha$ for any n.

We may also deduce 1.47 from 1.46, for $a_{n+p} > a_{n+1}$ for any n and $p > 1$, and so, taking a_{n+1} for λ in 1.46, we have $\alpha \geqslant a_{n+1} > a_n$.

1.5. If we wish to speak of the endless decimals x which lie between two decimals a and b, inclusive (where $a < b$), we say that *x lies in the interval* (a, b), or *x is a point in the interval* (a, b). The numbers a and b are called the end-points of the interval (a, b). Should it be necessary to exclude the end-points from the possible values of x we say that x lies in the *open interval* $[a, b]$, or if we wish to exclude only one end-point, suppose a, we say that x lies in the interval $[a, b)$ open on the left. To emphasize that an interval (a, b) includes its end-points we sometimes call (a, b) a *closed* interval. (Observe the use of the brackets '(', '[' to distinguish closed from open intervals.)

1.501. If c and d are both points in an interval (a, b) then we say that the interval (c, d) is *contained* in the interval (a, b), or that (c, d) is a *part of* (a, b).

1.6. Convergent series

A succession of sums, like $a_0, a_0+a_1, a_0+a_1+a_2, a_0+a_1+a_2+a_3,...$ is written for short in the form $a_0+a_1+a_2+a_3+...$ and is called an *infinite series.* If we denote a_0+a_1 by s_1, $a_0+a_1+a_2$ by s_2, $a_0+a_1+a_2+a_3$ by s_3, and so on, it follows that the infinite series $a_0+a_1+a_2+...$ stands for the sequence $s_1, s_2, s_3,...$. Accordingly we may speak of a convergent infinite series, the limit of an infinite series, the sum or product of two infinite series, and so on. The advantage of the series notation lies in the fact that, as we shall see, it serves to express the convergence condition in a rather simpler form. We shall frequently write $\sum_m^n a_r$ for

$$a_m+a_{m+1}+a_{m+2}+...+a_n$$

and $\sum a_r$ for the infinite series $a_0+a_1+a_2+...$, so that $\sum a_r$ stands for the sequence $\sum_0^1 a_r, \sum_0^2 a_r, \sum_0^3 a_r,...$.

1.601. The geometric series

The series $1+k+k^2+k^3+...$ is known as a *geometric series.* We shall show that the series is convergent for any value of k in the open interval $[-1, +1]$.

For any n, let s_n denote $1+k+k^2+...+k^n$. We can find two simple relations between s_n and s_{n+1}. First we notice that s_{n+1} is formed by adding k^{n+1} to s_n so that $s_{n+1} = s_n+k^{n+1}$, and secondly that if we multiply each term of s_n by k we obtain s_{n+1} apart from the initial unit, and therefore $s_{n+1} = ks_n+1$. Equating the two values of s_{n+1} we find

$$ks_n+1 = s_n+k^{n+1}$$

and so

$$s_n(1-k) = 1-k^{n+1} \quad \text{whence} \quad s_n = 1/(1-k)-k^{n+1}/(1-k),$$

provided $k \neq 1$.

Since $k^n \to 0$ when $-1 < k < 1$, and since the constant sequence $1-k, 1-k, 1-k,...$ has the limit $1-k$, it follows that

$$\lim s_n = 1/(1-k), \quad \text{provided } -1 < k < 1,$$

and therefore s_n is convergent for any value of k in the open interval $[-1, 1]$, and outside this interval the sequence s_n is not convergent since k^n is not convergent. When $k = -1$ we have already seen that k^n is not convergent and therefore s_n is not convergent. For the value $k = 1$, the equation

$$s_n = 1/(1-k)-k^{n+1}/(1-k)$$

does not hold, but for this value s_n is a sum of $n+1$ units and so $s_n = n+1$ and the sequence is not convergent.

1.6011. If $\sum a_n$ is convergent and if k does not depend on n, then $\sum ka_n$ is convergent.

For if $s_n = \sum\limits_1^n a_r$ and if $s_n \to s$ then $\sum\limits_1^n ka_r = ks_n \to ks$ and so $\sum ka_r$ is convergent.

1.6012. If $\sum a_n$ is convergent and if $\sum\limits_0^n a_r \to s$ then

$$(a_0+a_1)+(a_2+a_3)+(a_4+a_5)+\ldots$$

is convergent and its limit is s.

For $(a_0+a_1)+(a_2+a_3)+(a_4+a_5)+\ldots+(a_{2n}+a_{2n+1}) = \sum\limits_0^{2n+1} a_r \to s.$

Similarly $(a_0+a_1+\ldots+a_{r_1})+(a_{r_1+1}+\ldots+a_{r_2})+(a_{r_2+1}+\ldots+a_{r_3})+\ldots$ is convergent with limit s.

1.602. If $\sum a_n$ is convergent then the sequence a_n is convergent and $\lim a_n = 0$.

Since $\sum a_n$ is convergent we can find a sequence n_r such that

$$\sum_1^N a_n - \sum_1^{n_r} a_n = 0(r),$$

for any $N \geqslant n_r$, and so

$$a_N = \sum_1^N a_n - \sum_1^{N-1} a_n = \Big(\sum_1^N a_n - \sum_1^{n_r} a_n\Big) - \Big(\sum_1^{N-1} a_n - \sum_1^{n_r} a_n\Big)$$
$$= 0(r)-0(r) = 0(r-1),$$

provided $N > n_r$, which shows that $\lim a_n = 0$.

1.6021. The converse of 1.602 is not true for it can easily be shown that there are divergent series $\sum a_n$ for which $\lim a_n = 0$. Consider the series $1+\frac{1}{2}+\frac{1}{2}+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{4}+\frac{1}{4}+\ldots$, in which each of the $2n+1$ terms after the (n^2)th and up to, and including,

the $(n+1)^2$th equals $1/(2n+1)$. Here $\lim a_n = 0$, for $a_n = 0(r)$ if $n > 10^{2r}$, and $\sum a_n$ is divergent since

$$\sum_1^{(n+1)^2} a_r - \sum_1^{n} a_r > \sum_1^{(n+1)^2} a_r - \sum_1^{n^2} a_r = \{(n+1)^2-n^2\}/(2n+1) = 1.$$

However, 1.602 *gives a simple test for divergence, for it implies that if* a_n *is not convergent, or if* $\lim a_n$ *is other than zero, then* $\sum a_n$ *is divergent.*

1.61. Series of positive terms

If each term a_n is positive, $\sum a_n$ is called a *positive* series.

1.611. If $\sum a_n$ and $\sum b_n$ are positive series and if for some ν and h, $b_n \leqslant ha_n$, when $n \geqslant \nu$, then the convergence of $\sum a_n$ entails the convergence of $\sum b_n$.

For

$$\sum_1^N b_r - \sum_1^n b_r = b_{n+1}+b_{n+2}+...+b_N \leqslant h(a_{n+1}+a_{n+2}+...+a_N)$$
$$= \left\{\sum_1^N a_r - \sum_1^n a_r\right\}h,$$

provided $n \geqslant \nu$, and so, since $\sum_1^N a_r - \sum_1^{n_k} a_r = 0(k)$ for a certain n_k it follows that $\sum_1^N b_r - \sum_1^{n_k} b_r = h.0(k)$, which proves that $\sum b_r$ is convergent.

1.612. If $\sum a_n$ is a positive series and if $n_0, n_1, n_2, n_3,..$ is any sequence in which each term is less than its successor, then $\sum a_{n_r}$ is convergent, when $\sum a_n$ converges.

For n_r is greater than the r numbers $n_0, n_1, n_2,...$ which precede it and so $n_r > r$, whence, for $N > m$,

$$\sum_{r=1}^N a_{n_r} - \sum_{r=1}^m a_{n_r} = a_{n_{m+1}}+a_{n_{m+2}}+...+a_{n_N} < a_{m+1}+a_{m+2}+...+a_{n_N}.$$

But $a_{m+1}+a_{m+2}+...+a_{n_N} = \sum_1^{n_N} a_r - \sum_1^m a_r = 0(k)$ for a suitable value of m, and therefore $\sum_1^N a_{n_r} - \sum_1^m a_{n_r} = 0(k)$ which proves that $\sum a_{n_r}$ is convergent. $\sum a_{n_r}$ is called a *sub-series* of $\sum a_n$, so we may express our result by saying that a sub-series of a positive convergent series is convergent.

1.613. $\sum a_n$ is a positive series and the sequence a_{n+1}/a_n is convergent, with $\lim a_{n+1}/a_n = l$. Then if $l < 1$, $\sum a_n$ is convergent and if $l > 1$, $\sum a_n$ is divergent.

Let $m = |l-1|/2$ and choose k so that $m > 1/10^k$. Since $\lim a_{n+1}/a_n = l$ we can find a ν so that, for $n \geqslant \nu$, $a_{n+1}/a_n = l+0(k)$, and therefore $l-m < a_{n+1}/a_n < l+m$.

Suppose $l < 1$. Then $m = (1-l)/2$ and

$$a_{n+1}/a_n < l+(1-l)/2 = (l+1)/2 = r$$

say; since $(l+1)/2 < 1$ when $l < 1$ we have shown that for all $n \geqslant \nu$, $a_{n+1}/a_n < r < 1$. Thus $a_{\nu+1} < ra_\nu$, $a_{\nu+2} < ra_{\nu+1} < r^2a_\nu$, $a_{\nu+3} < ra_{\nu+2} < r^3a_\nu$, and so on. But $\sum r^n$ is convergent for $r < 1$ and $a_{\nu+p} < a_\nu r^p$, so that, by 1.611, $\sum a_{\nu+n}$ is convergent. Since $\sum_1^{N+\nu} a_r - \sum_1^{n+\nu} a_r = \sum_1^N a_{\nu+r} - \sum_1^n a_{\nu+r}$, the convergence of $\sum a_r$ follows from the convergence of $\sum a_{\nu+r}$.

Suppose next that $l > 1$. Then $m = (l-1)/2$ and

$$a_{n+1}/a_n > l-(l-1)/2 = (l+1)/2 > 1,$$

provided $n \geqslant \nu$. Thus

$$a_{\nu+1} > a_\nu, \quad a_{\nu+2} > a_{\nu+1} > a_\nu, \quad a_{\nu+3} > a_{\nu+2} > a_\nu,$$

and so on; i.e. $a_n > a_\nu$ for any $n > \nu$. Since $a_n > a_\nu$, a_n does not approach zero, and therefore $\sum a_n$ is divergent.

Observe that we draw no conclusion from the case $l = 1$.

1.7. Absolute convergence

$\sum a_n$ is said to be *absolutely convergent* if the positive series $\sum |a_n|$ is convergent.

An absolutely convergent series is also convergent.

For $\quad |a_{n+1}+a_{n+2}+...+a_N| \leqslant |a_{n+1}|+|a_{n+2}|+...+|a_N|$,

i.e. $$\left|\sum_1^N a_r - \sum_1^n a_r\right| \leqslant \sum_1^N |a_r| - \sum_1^n |a_r| = 0(k)$$

for a suitable value of n.

1.8. Power series

A series of the form $a_0+a_1x+a_2x^2+a_3x^3+...$ is called a *power series*.

1.81. If $|a_n/a_{n+1}|$ is convergent and $\lim |a_n/a_{n+1}| = R$ then the power series $\sum a_n x^n$ is absolutely convergent for any x in the *open*

interval $[-R, R]$, which is called the *interval of convergence* of the series.

This is an immediate consequence of 1.613, for if $|a_n/a_{n+1}| \to R$ then $|a_{n+1}/a_n| \to 1/R$, and therefore

$$|a_{n+1}x^{n+1}/a_n x^n| = |x||a_{n+1}/a_n| \to |x|/R,$$

so that $\sum |a_n x^n|$ is convergent if $|x| < R$, i.e. if x lies in the open interval $[-R, R]$. If $|x| > R$, 1.613 shows also that $\sum |a_n x^n|$ is divergent, but this does not suffice to prove that $\sum a_n x^n$ is divergent. We showed, however, in the second half of 1.613 that if $a_{n+1}/a_n > 1$ for $n \geqslant \nu$ then a_n does not approach zero, and therefore since $|a_{n+1}x^{n+1}|/|a_n x^n| \to |x|/R > 1$, when $|x| > R$, it follows that $a_n x^n$ does not tend to zero, which proves that $\sum a_n x^n$ is divergent.

Thus if $\lim |a_n/a_{n+1}| = R$ then the power series $\sum a_n x^n$ is absolutely convergent for any x in $[-R, R]$ and divergent for any x outside this interval. At either of the points $x = \pm R$ the series may converge or diverge.

1.82. If $\sum a_n X^n$ is convergent, then $\sum a_n x^n$ is absolutely convergent for any x such that $|x| < |X|$.

Since $\sum a_n X^n$ is convergent, $\lim a_n X^n = 0$, and therefore we can find N such that $|a_n X^n| < 1$ for $n \geqslant N$.

Thus $|a_n x^n| = |a_n X^n|.|(x/X)^n| < |x/X|^n$; but when $|x/X| < 1$, $\sum |x/X|^n$ is convergent and therefore $\sum |a_n x^n|$ is convergent, i.e. $\sum a_n x^n$ is absolutely convergent.

1.83. If $\sum a_n X^n$ is convergent, then $\sum na_n x^n$ is absolutely convergent for any x such that $|x| < |X|$.

First we notice that between $|X|$ and $|x|$, less than $|X|$, we can insert an $|x'|$ so that $|x| < |x'| < |X|$. Let $|x'|/|x| = 1+y$ and choose an $N > (2/y^2)+1$. Then

$$|x'/x|^n = (1+y)^n = 1+ny+\frac{n(n-1)}{2}y^2+... > \frac{n(n-1)}{2}y^2.$$

But if $n \geqslant N$, $(n-1)\frac{1}{2}y^2 \geqslant (N-1)\frac{1}{2}y^2 > 1$ so that $|x'/x|^n > n$ and therefore $n|x|^n < |x'|^n$. Since $|x'| < |X|$, $\sum a_n x'^n$ is absolutely convergent (by 1.82) and so, because $n|x|^n < |x'|^n$, $\sum na_n x^n$ is absolutely convergent for $|x| < |X|$ (by 1.611).

1.9.* $p_1, p_2, p_3, \dots$ is a steadily decreasing sequence of positive numbers and $b_1, b_2, b_3, \dots$ is *any* sequence the sum of the first n

terms of which is s_n. Each of the numbers $s_1, s_2, s_3, \dots, s_\kappa$ is less than h. Then

$$p_1 b_1 + p_2 b_2 + p_3 b_3 + \dots + p_\kappa b_\kappa < p_1 h.$$

Since $b_r = s_r - s_{r-1}$ for any r from 2 to κ and $s_1 = b_1$ we have

$$\begin{aligned} p_1 b_1 + p_2 b_2 + \dots + p_\kappa b_\kappa \\ &= p_1 s_1 + p_2(s_2 - s_1) + p_3(s_3 - s_2) + \dots + p_\kappa(s_\kappa - s_{\kappa-1}) \\ &= s_1(p_1 - p_2) + s_2(p_2 - p_3) + \dots + s_{\kappa-1}(p_{\kappa-1} - p_\kappa) + s_\kappa p_\kappa. \end{aligned}$$

As each p_r is less than p_{r-1} we may multiply the inequality $s_{r-1} < h$ by the positive factor $p_{r-1} - p_r$, giving

$$s_{r-1}(p_{r-1} - p_r) < h(p_{r-1} - p_r) \quad \text{for} \quad 1 < r \leqslant \kappa.$$

Furthermore p_κ is positive and therefore $s_\kappa p_\kappa < h p_\kappa$. Thus

$$\begin{aligned} s_1(p_1 - p_2) + s_2(p_2 - p_3) + \dots + s_{\kappa-1}(p_{\kappa-1} - p_\kappa) + s_\kappa p_\kappa \\ &< h(p_1 - p_2) + h(p_2 - p_3) + \dots + h(p_{\kappa-1} - p_\kappa) + h p_\kappa \\ &= h(p_1 - p_2 + p_2 - p_3 + \dots + p_{\kappa-1} - p_\kappa + p_\kappa) = h p_1, \end{aligned}$$

which completes the proof that $p_1 b_1 + p_2 b_2 + p_3 b_3 + \dots + p_\kappa b_\kappa < h p_1$.

Of course, if each b_r is positive the result is obvious, since

$$\begin{aligned} p_1 b_1 + p_2 b_2 + \dots + p_\kappa b_\kappa &< p_1 b_1 + p_1 b_2 + \dots + p_1 b_\kappa \\ &< p_1(b_1 + b_2 + \dots + b_\kappa) < p_1 h, \end{aligned}$$

but the proof we have given does not require that all b_r have the same sign, and the interest and importance of 1.9 lies just in this fact.

1.901.* Under the same conditions as 1.9, if each $|s_r|$ is less than H then

$$|p_1 b_1 + p_2 b_2 + \dots + p_\kappa b_\kappa| < p_1 H.$$

For

$$\begin{aligned} |s_1(p_1 - p_2) + s_2(p_2 - p_3) + \dots + s_{\kappa-1}(p_{\kappa-1} - p_\kappa) + s_\kappa p_\kappa| \\ &\leqslant |s_1|(p_1 - p_2) + |s_2|(p_2 - p_3) + \dots + |s_\kappa| p_\kappa < H p_1. \end{aligned}$$

1.91.* $\sum a_n$ is convergent and p_n is a positive decreasing sequence. Then $\sum p_n a_n$ is convergent.

We can choose n so that $|a_{n+1} + a_{n+2} + \dots + a_{n+\kappa}|$ is less than $1/10^r$ (for any κ). Hence, by 1.901,

$$|p_{n+1} a_{n+1} + p_{n+2} a_{n+2} + \dots + p_{n+\kappa} a_{n+\kappa}| < p_{n+1}/10^r < p_0/10^r,$$

which proves that $\sum p_n a_n$ is convergent.

1.911.* If $\sum a_n x^n$ is convergent then $\sum a_n \dfrac{x^{n+1}}{n+1}$ is convergent.

By 1.6011 the convergence of $\sum a_n x^{n+1}$ follows from that of $\sum a_n x^n$. Take $p_n = \dfrac{1}{n+1}$ in 1.91 and we see that $\sum a_n \dfrac{x^{n+1}}{n+1}$ is convergent.

1.912. If $\sum a_n x^n$ is absolutely convergent then $\sum a_n \dfrac{x^{n+1}}{n+1}$ is absolutely convergent; in this case we do not require 1.91, for $\sum a_n x^{n+1}$ is absolutely convergent and $\left| a_n \dfrac{x^{n+1}}{n+1} \right| < |a_n x^{n+1}|$ so that $\sum a_n \dfrac{x^{n+1}}{n+1}$ is absolutely convergent

1.913.* If $s_n = a_0 + a_1 + a_2 + \ldots + a_n$ and if each $|s_r|$ is less than H, and if p_n is a positive decreasing sequence such that $\lim p_n = 0$ then $\sum p_n a_n$ is convergent.

For by 1.901, $|p_{n+1} a_{n+1} + p_{n+2} a_{n+2} + \ldots + p_{n+\kappa} a_{n+\kappa}| < p_{n+1} H$, for any κ, and since $p_n \to 0$, therefore $p_n H \to 0$, and so we can choose n so that $p_{n+1} H = 0(r)$, which proves that $\sum p_n a_n$ is convergent.

The object of the next theorem is to show that throughout the interior of a closed interval in which it converges, a power series may be replaced by a polynomial, with a resulting error which may be made as small as we please.

1.92.* Let $s_n(x)$ denote $\sum\limits_1^n a_n x^n$ and $s(x) = \lim s_n(x)$ whenever the sequence $s_n(x)$ is convergent.

If $\sum a_n X^n$ is convergent then

$$s(x) = s_{n_r}(x) + 0(r)$$

for *any* x between 0 and X inclusive and for an n_r which depends *only* on r.

Since $\sum a_n X^n$ is convergent we can determine an n, a function of r, such that, for any κ,

$$|a_{n+1} X^{n+1} + a_{n+2} X^{n+2} + \ldots + a_{n+\kappa} X^{n+\kappa}| < 1/10^{r+1}.$$

Hence by 1.901, if $0 < x/X < 1$

$$\left| \left(\frac{x}{X}\right)^{n+1} a_{n+1} X^{n+1} + \left(\frac{x}{X}\right)^{n+2} a_{n+2} X^{n+2} + \ldots + \left(\frac{x}{X}\right)^{n+\kappa} a_{n+\kappa} X^{n+\kappa} \right|$$

$$< \left(\frac{x}{X}\right)^{n+1} \Big/ 10^{r+1} < 1/10^{r+1}$$

because $(x/X)^{n+1}$, $(x/X)^{n+2}$, $(x/X)^{n+3}$,... is steadily decreasing (and each is positive and less than unity), and so

$$|a_{n+1}x^{n+1}+a_{n+2}x^{n+2}+...+a_{n+\kappa}x^{n+\kappa}| < 1/10^{r+1},$$

whence $\quad s_{n+\kappa}(x)-s_n(x) = 0(r+1), \quad$ for $\quad 0 \leqslant x/X \leqslant 1.$

Since $s_n(x)$ is convergent and $s(x) = \lim s_n(x)$, *for each* x between 0 and X, we can find a κ (which is a function of x) so that

$$s(x)-s_{n+\kappa}(x) = 0(r+1)$$

and so $\quad s(x) = s_n(x)+0(r),$

where n depends on r but *not* on x.

The difficulty in this theorem lies in showing that we can find an $s_n(x)$ within a chosen distance of $s(x)$ without varying the number of terms of the series (i.e. without varying n) as x takes different values from 0 to X. If we choose an x to start with, of course we can find an n for which $s(x) = s_n(x)+0(r)$, for this value of x, simply because $\sum a_n x^n$ is convergent for this value of x, but this does not necessarily imply that the various values of n we consider have a greatest value. We might express 1.92 by saying that to find the values of $s(x)$ to a *chosen* degree of accuracy we need concern ourselves only with the values of a *fixed* $s_n(x)$.

1.921. If $\sum a_n X^n$ is absolutely convergent, the existence of n_r, depending only upon r, such that

$$s(x) = s_{n_r}(x)+0(r)$$

for any x satisfying $|x| \leqslant |X|$, follows without appeal to 1.901; for, if $|x| \leqslant |X|$,

$$\begin{aligned}|a_{n+1}x^{n+1}&+a_{n+2}x^{n+2}+...+a_{n+\kappa}x^{n+\kappa}|\\ &\leqslant |a_{n+1}||x|^{n+1}+...+|a_{n+\kappa}||x|^{n+\kappa}\\ &\leqslant |a_{n+1}||X|^{n+1}+|a_{n+2}||X|^{n+2}+...+|a_{n+\kappa}||X|^{n+\kappa},\end{aligned}$$

and since $\sum a_n X^n$ is absolutely convergent we can determine an n depending only upon r, such that, for any κ,

$$|a_{n+1}||X|^{n+1}+|a_{n+2}||X|^{n+2}+...+|a_{n+\kappa}||X|^{n+\kappa} = 0(r+1),$$

and the proof is completed as above; we have

$$s_{n+\kappa}(x)-s_n(x) = 0(r+1),$$

and since $s_n(x) \to s(x)$, for each value of x between 0 and X, we can find a κ (depending on x) so that $s(x)-s_{n+\kappa}(x) = 0(r+1)$, and therefore $s(x) = s_n(x)+0(r)$, where n depends only upon r.

1.93. The number $b-a$ is called the *length* of the interval (a, b); the length of the open interval $[a, b]$ is also $b-a$.

1.94. Consider the interval $(0, 1)$. Divide the interval into ten equal parts, that is to say, divide $(0, 1)$ into the ten intervals $(0, \cdot 1)$, $(\cdot 1, \cdot 2)$, $(\cdot 2, \cdot 3)$, and so on up to $(\cdot 9, 1)$. Choose one of these intervals, say $(\cdot 3, \cdot 4)$, and divide it into ten equal parts $(\cdot 30, \cdot 31)$, $(\cdot 31, \cdot 32)$, $(\cdot 32, \cdot 33)$,..., $(\cdot 39, \cdot 40)$ and select one of them, say $(\cdot 37, \cdot 38)$. Divide the chosen interval into ten equal parts, select one of them, say $(\cdot 371, \cdot 372)$; then divide $(\cdot 371, \cdot 372)$ into ten equal parts, and so on. We obtain a succession of intervals $(\cdot 3, \cdot 4)$, $(\cdot 37, \cdot 38)$, $(\cdot 371, \cdot 372)$,... each of which is a tenth part of its predecessor. If we denote the nth interval by $(a_n, a_n+1/10^r)$ then a_r is the value of a_n to r places for any $r < n$ (for instance $\cdot 37$ is the value of $\cdot 371$ to 2 places and any a_n, $n > 3$, commences with $\cdot 371$). Let α be the endless decimal whose nth digit is the nth digit of a_n (so that $\alpha = \cdot 371...$).

It is readily seen that α lies in each of the intervals $(a_n, a_n+1/10^n)$ for a_n is the value of α to n places so that $\alpha \geqslant a_n$, and $a_n+1/10^n$ is greater than a_{n+1}, i.e. greater than the value of α to $n+1$ places, and so $\alpha < a_n+1/10^n$. Since $0 < \alpha - a_n < 1/10^n$, $\lim a_n = \alpha$ and so also $\lim(a_n+1/10^n) = \alpha$.

Thus we have shown that a succession of intervals, in (0, 1), each of which is a tenth part of its predecessor, determines an endless decimal which lies in each of the intervals, and which is the limit of the sequences of left-hand and right-hand end-points of the intervals.

1.95. To extend the result of 1.94 to any interval (a, b) we consider the relation $t = (x-a)/(b-a)$. To any point x in (a, b) corresponds a unique value of t in $(0, 1)$, for when $x = a$, $t = 0$ and when $x = b$, $t = 1$, and when x lies between a and b, t lies between 0 and 1, since $(x-a)/(b-a)$ steadily increases with x. Since $x = a+(b-a)t$, to a point t in $(0, 1)$ corresponds similarly a unique x in (a, b). Moreover, if (α, β) and (γ, δ) are two intervals in (a, b), of equal length, so that $\beta-\alpha = \delta-\gamma$, then the intervals

$$((\alpha-a)/(b-a),\ (\beta-a)/(b-a))$$

and

$$((\gamma-a)/(b-a),\ (\delta-a)/(b-a))$$

are of equal length, for

$$(\beta-a)/(b-a)-(\alpha-a)/(b-a) = (\beta-\alpha)/(b-a)$$
$$= (\delta-a)/(b-a)-(\gamma-a)/(b-a);$$

thus to equal intervals in (a, b) correspond equal intervals in $(0, 1)$, and vice versa. Hence if (a, b), or any part of (a, b), is divided into ten equal parts, so too $(0, 1)$, or a part of $(0, 1)$, is divided into ten equal parts, by the values of t corresponding to the values of x in (a, b). Finally we observe that if (γ, δ) is a part of (α, β), which is itself a part of (a, b), then the corresponding intervals

$$((\gamma-a)/(b-a), (\delta-a)/(b-a)), ((\alpha-a)/(b-a), (\beta-a)/(b-a)), (0, 1)$$

stand in the same relationship, since

$$0 \leqslant (\alpha-a)/(b-a) \leqslant (\gamma-a)/(b-a),$$

and
$$(\delta-a)/(b-a) \leqslant (\beta-a)/(b-a) \leqslant 1.$$

Thus if (a, b) is divided repeatedly into ten equal parts, determining a succession of intervals (a, b), (a_1, b_1), (a_2, b_2), (a_3, b_3),..., each of which is a tenth part of its predecessor, and if, for each n, $\lambda_n = (a_n-a)/(b-a)$, $\mu_n = (b_n-a)/(b-a)$ then $(0, 1)$, (λ_1, μ_1), (λ_2, μ_2), (λ_3, μ_3),... is a succession of intervals each of which is a tenth part of its predecessor, and so, by 1.94, there is a unique point ν which is the limit of both the sequences $\lambda_1, \lambda_2, \lambda_3,...$ and $\mu_1, \mu_2, \mu_3,...$ and which lies in each of the intervals (λ_n, μ_n). Let $X = a+(b-a)\nu$, then we shall show that X is the limit of both the sequences $a_1, a_2, a_3,...$ and $b_1, b_2, b_3,...$ and that X lies in all the intervals (a_n, b_n); for $\lambda_n \leqslant \nu \leqslant \mu_n$ and so

$$a_n = a+(b-a)\lambda_n \leqslant a+(b-a)\nu = X \leqslant a+(b-a)\mu_n = b_n,$$

i.e. $a_n \leqslant X \leqslant b_n$, and furthermore

$$X-a_n = a+(b-a)\nu-\{a+(b-a)\lambda_n\} = (b-a)(\nu-\lambda_n) \to 0,$$

i.e. $a_n \to X$, and similarly $b_n \to X$.

1.96. Theorem 1.94 may readily be generalized. If we divide (a, b) into any number of equal parts (two or more), choose one of the parts (a_1, b_1), divide (a_1, b_1) into any number of equal parts and choose one part (a_2, b_2) say, divide up (a_2, b_2), and so on, then there is one and only one point which lies in all the chosen intervals and this point is the limit of both the sequences (a_n) and (b_n).

For $b_n-a_n \leqslant \frac{1}{2}(b_{n-1}-a_{n-1})$ and so $b_n-a_n \leqslant \frac{1}{2^n}(b-a)$; moreover for $N > n$, a_N and b_N lie in (a_n, b_n) and so

$$0 \leqslant a_N-a_n \leqslant b_n-a_n \leqslant \frac{1}{2^n}(b-a)$$

and
$$0 \leqslant b_n-b_N \leqslant b_n-a_n \leqslant \frac{1}{2^n}(b-a)$$

Hence, since $\frac{1}{2^n}(b-a)$ may be made as small as we please by choosing n great enough, both (a_n) and (b_n) are convergent and $b_n-a_n \to 0$. Thus if l and l' are the limits of (a_n) and (b_n), from $l-a_n \to 0$, $l'-b_n \to 0$ and $b_n-a_n \to 0$ we deduce $l-l' = 0(r)$ for any r and so $l = l'$. Furthermore, since $a_{n+1} \geqslant a_n$ therefore $l \geqslant a_n$, and since $b_{n+1} \leqslant b_n$ therefore $l = l' \leqslant b_n$, i.e. l lies in the interval (a_n, b_n), and this is true for any n.

II

CONTINUITY

Expressions like n^2+1, $\sqrt[3]{m}$, $\sqrt[5]{(n^2+1)}/\sqrt[3]{n}$ are called numeral *functions* (of one variable) if the letter which each contains may be replaced by any numeral; such as $(m+n)/(m^2+n^2)$, $\sqrt[3]{(m^7+n)}$, or $\sqrt[m]{n}$ are numeral functions of two variables, and so on. The letters in a functional expression are called the *arguments* of the function. The essential property of a functional expression (in the sense in which we use this term throughout this book) is that when the arguments are replaced by numbers the expression can be worked out, determining a unique number which is called the *value* of the function for the set of numbers which have replaced the arguments. For instance, since $(3+8)/(3^2+8^2) = 11/73$, the value of the function $(m+n)/(m^2+n^2)$, when $m = 3$ and $n = 8$, is $11/73$. Functions of a single variable are sometimes called sequences; in fact the only difference between a function and a sequence is that in a sequence we think of the values of the function as taken in a definite order, the value of the function when the argument is replaced by 1 forming the first term of the sequence, the value when the argument is replaced by 2 forming the second term, and so on.

Expressions in which the letters may be replaced by endless decimals (or terminating decimals) are known as decimal functions, or, shortly, as functions. Decimal functions are of many kinds. In addition to the familiar expressions of elementary algebra, like x^2+5x+7 or $\sqrt[5]{\{x^2/(x^3+x-1)\}}$, any convergent series

$$a_0+a_1x+a_2x^2+\dots$$

is a function, the limit of the series for a given value of x being the value of the function. From any two functions another can be formed by replacing the argument of one of the functions by the other function; thus from the functions $x^{\frac{1}{2}}$ and x^5+x+1 we can form the function $(x^5+x+1)^{\frac{1}{2}}$ or the function $x^{\frac{5}{2}}+x^{\frac{1}{2}}+1$.

Decimal functions are commonly classified under the headings 'Polynomial', 'Rational', 'Algebraic', and 'Transcendental'. A polynomial is a terminating series of the form

$$a_0+a_1x+a_2x^2+\dots+a_nx^n,$$

a rational function is a fraction in which both the numerator and denominator are polynomials, and an expression formed by combinations of fractional powers of rational functions is an algebraic function; an endless convergent series which has neither a rational nor an algebraic limit is a transcendental function.

Though the functions most commonly met with in the calculus are of the kinds we have described we have by no means exhausted the unlimited variety of function definitions that play one part or another in mathematics. We might for instance define a numeral function by saying that the value of the function is zero for an even value of the argument and unity for an odd value; for such a function there is no question of *calculating* its *value*, but of calculating whether the value of the argument has a certain property, or not, in this case whether it be even or odd. Other definitions of this kind are 'the whole part of x', a function which has a constant value for any value of the argument between consecutive integers, and the function which is defined to have the value zero if the endless decimal x contains the digit 3 and the value unity if x does not contain the digit 3.

A function may be defined for any value of its argument or only for certain selected values. The function $\sqrt{\{(x-1)(2-x)\}}$, for instance, is defined only for values of x between 1 and 2, since a negative number has no square root. We shall denote an unspecified numeral function of one, two, three variables, etc., by $f(n)$, $f(m, n)$, $f(l, m, n)$, etc., respectively, and unspecified decimal functions by $f(x)$, $f(x, y)$, $f(x, y, z)$, etc., where f may be replaced by any other single letter.

2. Continuity

Roughly speaking, a continuous function is one which changes its value but little for a small change in the value of its argument. If we think of the function as evaluated to a certain number of decimal places, say to n places, then for the function to be continuous we require that its value to n places be unchanged by any sufficiently small change in its argument. For example, the function $\sqrt{x}$, worked out to 2 decimal places, has the constant value ·35 for any x between ·1225 and ·1296, or if worked out to 3 places, the constant value ·352 for any x between ·123904 and ·124609.

2.01. Formally we define: $f(x)$ is uniformly (or interval) continuous in an interval (a, b) if we can find a function $n(r)$ such that, for any X and x in (a, b),

$$f(X)-f(x) = 0(r) \quad \text{provided } X-x = 0(n(r)).$$

For brevity we shall generally write 'continuous' for 'uniformly continuous'.

2.02. EXAMPLE. x^2 is a continuous function in any interval (a, b). Consider X^2-x^2. If X and x are both between a and b then $X+x$ lies between $2a$ and $2b$ so that if h is the greater of the two numbers $2|a|$, $2|b|$ it follows that $|X+x| \leqslant h$. Choose k so that $h < 10^k$ (k is the number of digits in the whole part of h). Hence if

$$X-x = 0(r+k)$$

we have

$$|X^2-x^2| = |X-x||X+x| < 10^k/10^{r+k} = 1/10^r.$$

Thus for any X, x in (a, b),

$$X^2-x^2 = 0(r) \quad \text{provided } X-x = 0(r+k)$$

(where k depends only upon a and b), so that x^2 is continuous.

2.03. A similar argument shows that x^n is continuous, for any numeral n. For if c is the greater of the two $|a|$, $|b|$ then, for any p between 1 and n, $|X^{n-p}x^{p-1}| < c^{n-p}c^{p-1} = c^{n-1}$ and therefore $X^{n-1}+X^{n-2}x+X^{n-3}x^2+...+x^{n-1}$ has an absolute value less than nc^{n-1}. Hence if we choose k so that $10^k > nc^{n-1}$ and if

$$X-x = 0(r+k)$$

we have

$$|X^n-x^n|$$
$$= |X-x||X^{n-1}+X^{n-2}x+X^{n-3}x^2+...+x^{n-1}| < 10^k/10^{k+r} = 1/10^r$$

and so $X^n-x^n = 0(r)$, which proves that x^n is continuous in any interval.

2.04. The function $1/x$ is continuous in the interval (a, b) provided the interval does not contain the value zero (i.e. provided $ab > 0$). For if x and X are both between a and b, and if c is the lesser of $|a|$, $|b|$ then $|xX| > c^2 > 0$ and therefore

$$|1/X-1/x| = |X-x|/|xX| < |X-x|/c^2.$$

Since $c^2 > 0$ we can choose k so that $c^2 \geqslant 1/10^k$ (the kth is the first non-zero decimal digit of c^2, if $c^2 < 1$, otherwise $k = 0$) and so

$$1/X-1/x = 0(r) \quad \text{provided } X-x = 0(r+k).$$

2.1. ***A function which is continuous in an interval (a, b) is bounded in that interval***, i.e. we can find numbers m, M between which all the values of the function lie.

If $f(x)$ is continuous in (a, b) we can find $n(r)$ so that for any x, X in (a, b)

$$f(X)-f(x) = 0(r) \quad \text{provided } X-x = 0(n(r)).$$

Choose k so that $b-a < 10^k$ and divide the interval (a, b) into $10^{k+n(1)}$ equal parts so that the length of each part is less than $1/10^{n(1)}$. Let $a_1, a_2, a_3,\ldots, a_p$ be the points of subdivision. Since the length of each of the intervals (a_r, a_{r+1}) is less than $1/10^{n(1)}$, the continuity of $f(x)$ shows that if X is any point in the interval (a_r, a_{r+1}) then $f(X)-f(a_r) = 0(1)$. Now any point X in the interval (a, b) must fall into one of the parts into which we have divided the interval, so we may suppose it lies in the part (a_r, a_{r+1}). Adding the equations $f(a_1)-f(a) = 0(1)$, $f(a_2)-f(a_1) = 0(1)$, $f(a_3)-f(a_2) = 0(1),\ldots, f(a_r)-f(a_{r-1}) = 0(1)$, $f(X)-f(a_r) = 0(1)$ we find $f(X)-f(a) = (r+1).0(1)$ and therefore

$$|f(X)-f(a)| < (r+1)/10 \leqslant (p+1)/10,$$

where p is the number of points of subdivision. Since the interval is divided into $10^{k+n(1)}$ parts, $p+1 = 10^{k+n(1)}$ and therefore for any X in (a, b), $|f(X)-f(a)| < 10^{k+n(1)-1}$ so that $f(X)$ lies between the numbers $f(a)-10^{k+n(1)-1}$ and $f(a)+10^{k+n(1)-1}$.

2.2. If $f(x)$ is continuous in (a, b) and if (c, d) is a part of (a, b) then $f(x)$ is continuous in (c, d).

$f(X)-f(x) = 0(r)$ provided $X-x = 0(n(r))$ for any X, x in (a, b) and so for any X, x in (c, d).

2.21. If $f(x)$ is continuous in (a, b) and continuous in (b, c) then $f(x)$ is continuous in (a, c).

We can find $n(r)$ so that

$$f(X)-f(x) = 0(r+1) \quad \text{for} \quad X-x = 0(n(r))$$

provided x, X *both* lie in (a, b) or *both* in (b, c).

If x lies in (a, b) and X in (b, c) then

$$f(X)-f(b) = 0(r+1),$$
$$f(x)-f(b) = 0(r+1)$$

provided $X-b$ and $x-b$ are $0(n(r))$.

Thus, for any X, x in (a, c),

$$f(X)-f(x) = 0(r) \quad \text{provided } X-x = 0(n(r)).$$

2.22. If (a, b) can be divided into $p(r)$ parts (of which the smallest has a length $l(r)$) such that for any x, X in the same part

$$f(X)-f(x) = 0(r)$$

then $f(x)$ is continuous in (a, b); for if $|X-x| < l(r)$ then x and X lie either in the same part, in which case $f(X)-f(x) = 0(r)$, or they lie in adjacent parts, separated by some point c common to both parts so that $f(x)-f(c) = 0(r)$ and $f(X)-f(c) = 0(r)$ and therefore

$$f(X)-f(x) = 0(r-1).$$

2.23. If $f(x)$ is steadily increasing in (a, b) and if each number between $f(a)$ and $f(b)$ is a value of $f(x)$ then $f(x)$ is continuous in (a, b).

Divide $(f(a), f(b))$ into n equal parts, choosing n so that

$$\{f(b)-f(a)\}/n = 0(r).$$

Let $f(a) = y_0, y_1, y_2,..., y_n = f(b)$ be the points of subdivision. For each r, $f(a) \leqslant y_r \leqslant f(b)$ and so there is a point x_r such that $f(x_r) = y_r$. Since $f(x)$ is increasing, and $y_{r+1} > y_r$, therefore $x_{r+1} > x_r$, so that $a = x_0, x_1, x_2,..., x_n = b$ is a subdivision of (a, b). If $x_p \leqslant x < X \leqslant x_{p+1}$, then

$$f(X)-f(x) \leqslant f(x_{p+1})-f(x_p) = y_{p+1}-y_p = \{f(b)-f(a)\}/n = 0(r)$$

and therefore, by 2.22, $f(x)$ is continuous in (a, b).

2.3. If $f(x)$ is continuous in (a, b) and if $c_1, c_2, c_3,...$ is a convergent sequence of points in (a, b) which has a limit c also in (a, b) then the sequence $f(c_1), f(c_2), f(c_3),...$ is convergent and $f(c)$ is its limit.

Since $f(x)$ is continuous we can find $n(r)$ so that for any X and x in (a, b)

$$f(X)-f(x) = 0(r) \quad \text{provided } X-x = 0(n(r)).$$

Moreover since $c = \lim c_n$ we can find $k(r)$ so that

$$c-c_n = 0(r) \quad \text{for } n \geqslant k(r)$$

and therefore $\quad c-c_n = 0(n(r)) \quad \text{if } n \geqslant k(n(r)).$

Whence $\quad f(c)-f(c_n) = 0(r) \quad \text{for } n \geqslant k(n(r))$

which proves that $f(c_n) \to f(c)$.

This result may be expressed in the form

$$\lim f(c_n) = f(\lim c_n).$$

Theorem 2.3 plays an important part in the determination of the values of functions for *endless* decimal arguments. Suppose we wished to find the value of the function $\sqrt{x}$ for $x = \cdot 333\ldots$. The function $\sqrt{x}$ is continuous and therefore our theorem tells us that $\sqrt{\cdot 333\ldots}$ is the limit of the (necessarily) convergent sequence $\sqrt{\cdot 3}$, $\sqrt{\cdot 33}$, $\sqrt{\cdot 333}, \ldots$. Since we can evaluate the square root of a *terminating* decimal to as many decimal places as we please we can determine the limit of this sequence to as many places as we please and thus the value of $\sqrt{\cdot 333\ldots}$ is determined.

Since $\sqrt{\cdot 3} > \sqrt{\cdot 25} = \cdot 5$, $\sqrt{\cdot 333\ldots} - \sqrt{\cdot 33} = \dfrac{\cdot 333\ldots - \cdot 33}{\sqrt{\cdot 333\ldots}+\sqrt{\cdot 33}} < \cdot 00\dot{3}$

and therefore, as $\sqrt{\cdot 33} = \cdot 5744\ldots$, the value of $\sqrt{\cdot \dot{3}}$ is $\cdot 57$ to 2 decimal places, and the error in taking $\cdot 57$ for $\sqrt{\cdot \dot{3}}$ is less than $\cdot 0044\ldots + \cdot 00\dot{3}$ and so less than $\cdot 008$.

2.4. If $f(x)$ is continuous in (a, b) and if $f(a)$ is negative and $f(b)$ is positive then we can find a point between a and b where $f(x)$ takes the value zero.

First we show that if a and b are endless decimals we can find terminating decimals α and β, forming an interval (α, β) contained in the interval (a, b) and such that $f(\alpha)$ is negative and $f(\beta)$ is positive. Let α' and β' be the values of a and b to k places of decimals, and let $\alpha = \alpha' + 1/10^k$ and $\beta = \beta' - 1/10^k$ so that, provided k is chosen great enough, $a < \alpha < \beta < b$. Since each of the differences $a - \alpha$ and $b - \beta$ is $0(k)$, the differences $|f(a) - f(\alpha)|$ and $|f(b) - f(\beta)|$ can be made as small as we please by choosing k sufficiently large, for $f(x)$ is a continuous function, and so can be made less than both $\frac{1}{2}|f(a)|$ and $\frac{1}{2}|f(b)|$. Thus $f(\alpha)$ differs from the positive $f(a)$ by less than $\frac{1}{2}f(a)$ and is therefore also positive, and $f(\beta)$ differs from the negative $f(b)$ by less than $\frac{1}{2}|f(b)|$ and so $f(\beta)$ is negative.

Divide the interval (α, β) into $(\beta - \alpha)10^k$ equal parts (so that each part is of length $1/10^k$) and consider the value of $f(x)$ at each point of subdivision in turn, commencing at α. We obtain a succession of values, the first of which is negative and the last positive. As we pass along the successive values of $f(x)$ we shall reach a stage where the value of $f(x)$ *first* changes to a *positive* number—let

b_1 be the point of subdivision where this change occurs. If a_1 is the point of subdivision which immediately precedes b_1, $f(a_1)$ is less than (or equal to) zero, for b_1 is the first point of subdivision at which $f(x)$ is greater than zero. Next divide the interval (a_1, b_1) into 10 equal parts (each part will be of length $1/10^{k+1}$) and as before consider the values of $f(x)$ at each point of subdivision in turn from a_1 to b_1. As $f(a_1) \leqslant 0$ and $f(b_1) > 0$ we shall reach a first point of subdivision at which $f(x) > 0$ and let this point be called b_2, so that, if a_2 is the point of subdivision which immediately precedes b_2, $f(a_2) \leqslant 0$. Divide (a_2, b_2) into 10 equal parts (of length $1/10^{k+2}$) and let b_3 be the first point of subdivision at which $f(x) > 0$ and a_3 the point before b_3, and then divide (a_3, b_3) into 10 equal parts, and so on. By Theorem 1.95 the sequence a_1, a_2, a_3,... is convergent and tends to a limit c say, contained in all the intervals (a_r, b_r). Since $f(x)$ is continuous we can find $n(r)$ so that

$$f(X)-f(x) = 0(r)$$

provided $X-x = 0(n(r))$. Hence, as $b_n-a_n = 0(k+n-1)$,

$$f(b_n)-f(a_n) = 0(r) \quad \text{for} \quad n \geqslant n(r)-k+1.$$

Now $f(b_n) > 0$, $f(a_n) \leqslant 0$ and so $f(b_n)-f(a_n)$ is greater than $-f(a_n)$, which is positive, whence $-f(a_n) = 0(r)$, i.e. $f(a_n) = 0(r)$, for $n \geqslant n(r)-k+1$, so that $\lim f(a_n) = 0$. But $f(x)$ is continuous and $\lim(a_n) = c$ and therefore

$$f(c) = f(\lim a_n) = \lim f(a_n) = 0.$$

Thus we have found a point c between a and b such that $f(c) = 0$.

2.41. From 2.4 we deduce:

If $F(x)$ is continuous in (a, b) then, as x varies from a to b, $F(x)$ passes through every value between $F(a)$ and $F(b)$.

Suppose for instance that $F(a) < F(b)$ and let v be any number between $F(a)$ and $F(b)$. Denote the function $F(x)-v$ by $f(x)$. $f(x)$ is continuous as $f(X)-f(x) = F(X)-F(x)$ and

$$f(a) = F(a)-v < 0, \qquad f(b) = F(b)-v > 0.$$

Accordingly we can find a point c in the interval (a, b) such that $f(c) = 0$. But $F(c) = f(c)+v$ and so $F(c) = v$.

2.5. The sum and the difference of two continuous functions are continuous functions.

For $\{f(X)\pm g(X)\}-\{f(x)\pm g(x)\}=\{f(X)-f(x)\}\pm\{g(X)-g(x)\}$ and so if both $f(X)-f(x)$ and $g(X)-g(x)$ are $0(r)$, and if

$$h(x)=f(x)\pm g(x), \quad \text{then} \quad h(X)-h(x)=0(r-1)$$

and therefore $h(x)$ is continuous.

2.51. The product of continuous functions is a continuous function.

For if $f(X)-f(x)=0(r)$ and $g(X)-g(x)=0(r)$ then

$$\begin{aligned} f(X)g(X) &= \{f(x)+0(r)\}\{g(x)+0(r)\} \\ &= f(x)g(x)+\{f(x)+g(x)\}.0(r)+0(2r). \end{aligned}$$

Now $f(x)$ and $g(x)$, being continuous, are bounded and so we can find k so that $|f(x)+g(x)|<10^k$. Therefore

$$|f(X)g(X)-f(x)g(x)|<10^k/10^r+1/10^{2r}<1/10^{r-k-1}$$

which shows that $f(x)g(x)$ is continuous.

2.52. If $g(x)$ is continuous in (a, b) and if for some k, $|g(x)|>1/10^k$ for any x in (a, b), then $1/g(x)$ is continuous in (a, b).

We can choose $n(r)$ so that $g(X)-g(x)=0(r)$ for any x, X in (a, b) such that $X-x=0(n(r))$ and therefore

$$\left|\frac{1}{g(X)}-\frac{1}{g(x)}\right|=\left|\frac{g(x)-g(X)}{g(x)g(X)}\right|<\frac{1}{10^r}\Big/\frac{1}{10^{2k}}=\frac{1}{10^{r-2k}}.$$

Whence it follows that $\left|\frac{1}{g(X)}-\frac{1}{g(x)}\right|<\frac{1}{10^r}$, provided

$$X-x=0(n(r+2k)),$$

and so $1/g(x)$ is continuous.

2.53. If $f(x)$ is continuous in (a, b) and if $F(x)$ is continuous in an interval (α, β) which contains $f(x)$ for any x in (a, b), then $F(f(x))$ is continuous in (a, b).

For any Y, y in (α, β) we can choose s so that

$$F(Y)-F(y)=0(r) \tag{i}$$

provided $Y-y=0(s)$.

Furthermore for any X, x in (a, b) we can choose t so that

$$f(X)-f(x)=0(s) \quad \text{provided } X-x=0(t).$$

Hence, in equation (i), we may take $f(X)$ and $f(x)$ for Y and y giving

$$F(f(X))-F(f(x))=0(r), \quad \text{provided } X-x=0(t),$$

and therefore $F(f(x))$ is continuous in (a, b).

2.54. If $g(x)$ and $f(g(x))$ are continuous in (a, b), and if α, β are any two points in (a, b), then $f(t)$ is continuous in the interval $g(\alpha), g(\beta)$.

Proof. We may suppose α, β so chosen that $g(\alpha) < g(\beta)$.

Since $f(g(x))$ is continuous we can divide (α, β) into n equal parts by the points $\alpha = \alpha(0),\ \alpha(1),\ \alpha(2),\ldots,\ \alpha(n) = \beta$, say, such that $f(g(X)) - f(g(x)) = 0(k)$ for any two x, X in the same part.

Since $g(\alpha(n)) > g(\alpha(0))$, we can determine a least integer r_1 such that $g(\alpha(0)) < g(\alpha(r_1)) \leqslant g(\alpha(n))$, where r_1 may equal n; if $g(\alpha(r_1)) < g(\alpha(n))$ then we can determine a *least* r_2, greater than r_1, such that

$$g(\alpha(r_1)) < g(\alpha(r_2)) \leqslant g(\alpha(n)),$$

where r_2 may equal n, and so on, up to some r_m such that

$$g(\alpha(r_{m-1})) < g(\alpha(r_m)) = g(\alpha(n)).$$

For any $p \leqslant m$, since r_p is the least integer, greater than r_{p-1}, such that $g(\alpha(r_p)) > g(\alpha(r_{p-1}))$, therefore $g(\alpha(r_p - 1)) \leqslant g(\alpha(r_{p-1}))$, and so each interval $g(\alpha(r_{p-1})), g(\alpha(r_p))$ is contained in an interval $g(\alpha(r_p - 1)), g(\alpha(r_p))$, that is, in an interval bounded by values of $g(x)$ at *consecutive* points of the original subdivision of (α, β).

The interval $g(\alpha), g(\beta)$ is divided into m parts by the points $g(\alpha(r_q))$, $0 \leqslant q \leqslant m$, $r_0 = 0$; if t, T are any two points in the same part then both t and T lie between some pair $g(\alpha(i)), g(\alpha(i+1))$ so that, since $g(x)$ is continuous, we can determine x, X between $\alpha(i)$ and $\alpha(i+1)$ such that $g(x) = t$, $g(X) = T$ and therefore

$$f(T) - f(t) = f(g(X)) - f(g(x)) = 0(k),$$

which proves that $f(t)$ is continuous in the interval $g(\alpha), g(\beta)$.

2.6. If $\sum a_n X^n$ is convergent then $\sum a_n x^n$ is a continuous function in $(0, X)$.

Let $s_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$ and $s(x) = \lim s_n(x)$. By Theorem 1.92 we can find an m so that for any x in $(0, X)$

$$s(x) = s_m(x) + 0(r).$$

Since x^n is a continuous function (for any n), $s_m(x)$ is a sum of continuous functions and is therefore itself continuous, and so for a certain k,

$$s_m(X) - s_m(x) = 0(r) \quad \text{provided } X - x = 0(k).$$

Hence $s(X) - s(x) = 0(r) + 0(r) - 0(r) = 0(r - 1)$ and so $s(x)$ is continuous in $(0, X)$.

2.7. If $f(x)$ and $g(x)$ are continuous in some interval which includes a point X, and if for *all* x in this interval, *except* $x = X$, we have $f(x) = g(x)$, then also $f(X) = g(X)$.

For if X is not the right-end point of the interval then

$$f\left(X+\frac{1}{n}\right) = g\left(X+\frac{1}{n}\right),$$

for sufficiently great values of n, and by 2.3

$$f\left(X+\frac{1}{n}\right) \to f(X), \qquad g\left(X+\frac{1}{n}\right) \to g(X),$$

which proves that $f(X) = g(X)$; and if X is the right-end point then $X-1/n$ belongs to the interval for sufficiently great values of n and the proof is completed as before.

III

THE DERIVED FUNCTION

IMPLICIT FUNCTIONS. MEAN SLOPE. DERIVATIVE OF A POWER SERIES. THE BINOMIAL SERIES. INVERSE FUNCTIONS

THE object of the differential calculus is to compare with a change in the argument of a function the corresponding change in the value of the function, or, as we may say, to measure the *rate of change* of a function, that is the average change in value per unit change of argument. For a change in argument from x to X the value changes from $f(x)$ to $f(X)$, and the ratio of the amount by which the value changes to the amount of change in the argument is

$$\frac{f(X)-f(x)}{X-x}.$$

If, for each value of x, we can determine a number $f'(x)$, say, such that $\dfrac{f(X)-f(x)}{X-x}$ differs from $f'(x)$ by as little as we please for all values of X *sufficiently near* to x, the function $f(x)$ is said to be *regular* or *differentiable*, and $f'(x)$ is called the *derivative* of $f(x)$. $f'(x)$ records the rate at which $f(x)$ is changing its value at the point x. For instance, since $\dfrac{X^2-x^2}{X-x} = X+x$ and since $X+x$ is nearly equal to $x+x$, i.e. $2x$, when X is near x, it follows that $2x$ is the derivative of x^2. Thus the greater the value of x, the faster does x^2 change its value; at the point $x = 5$ the function x^2 is increasing 10 times as fast as its argument and at $x = 20$, x^2 is increasing 40 times as fast as its argument. Observe that in considering the ratio $\dfrac{X^2-x^2}{X-x}$ we are concerned with values of X *close* to x but *different from* x; when we say that the derivative of x^2 is $2x$ because $X+x$ is nearly equal to $2x$, the fact that $X+x$ actually equals $2x$ when X equals x is quite irrelevant, because the *only* values of X which we are considering are those near x and different from it, and $\dfrac{X^2-x^2}{X-x}$ is equal to $X+x$ only so long as X and x are unequal.

We do not attempt to measure the rate of change of the value of functions $f(x)$ for which $\dfrac{f(X)-f(x)}{X-x}$ varies appreciably for small variations in X.

Another way in which we might express the condition that a function be regular is to say that for small enough changes in the argument, changes in the value of the function are nearly proportional to changes in the argument, i.e. if X_1, X_2 are points near x, then $\dfrac{f(X_1)-f(x)}{f(X_2)-f(x)}$ is nearly equal to $\dfrac{X_1-x}{X_2-x}$. The difference columns in a mathematical table are constructed on this principle. For instance, in a four-figure table of squares we find that the squares of the numbers 1·52, 1·53, and so on up to 1·58 (each of which differs from its successor by ·01) are 2·310, 2·341, 2·372, 2·403, 2·434, 2·465, and 2·496. The difference between each of these and its successor is readily seen to be ·031, and so to equal changes in the argument correspond changes in the value of the square which coincide to 3 decimal places. If we wish to find the square of a number between any two of those we have just considered, say 1·543, which lies between 1·54 and 1·55, we observe that the difference between 1·54 and 1·543, i.e. ·003, is $\frac{3}{10}$ of the difference between 1·54 and 1·55. Hence the difference between the squares of 1·543 and 1·54 is about $\frac{3}{10}$ of the difference between the squares of 1·55 and 1·54. Since $1{\cdot}55^2-1{\cdot}54^2 = {\cdot}031$ it follows that $1{\cdot}543^2-1{\cdot}54^2$ must equal $\frac{3}{10}\times{\cdot}031 = {\cdot}009$, to 3 places. Adding this difference to $1{\cdot}54^2$ we find $1{\cdot}543^2 = 2{\cdot}381$; similarly the addition of the same ·009 to the squares of 1·52, 1·53, 1·55, etc., gives the squares of 1·523, 1·533, 1·553, etc., and 9 is the figure entered in the difference column for 3, in the row commencing 1·50. The justification of this method lies in the fact that x^2 is a regular function.

The reader who is familiar with coordinate geometry will recognize that the condition for a function $f(x)$ to be regular is just that any sufficiently small part of the curve $y = f(x)$ is nearly a straight line. Any three points (x, y), (x_1, y_1), (x_2, y_2) on a straight line satisfy the condition $\dfrac{y_1-y}{y_2-y} = \dfrac{x_1-x}{x_2-x}$; for a differentiable func-

tion $f(x)$, if the points lie on $y = f(x)$, the ratio $\frac{y_1 - y}{y_2 - y}$ is *nearly* equal to $\frac{x_1 - x}{x_2 - x}$, provided x_1 and x_2 are both close to x.

3. Formal definition

A function $f(x)$ is said to be uniformly (or interval) differentiable in an interval (a, b) if we can find functions $f'(x)$ and $n(r)$ such that

$$\frac{f(X)-f(x)}{X-x} = f'(x)+0(r) \tag{D}$$

for any x, X in (a, b) satisfying $X-x = 0(n(r))$.

The function $f'(x)$ is called the uniform (or interval) derivative of $f(x)$, in (a, b). For brevity we shall generally write 'differentiable' for 'uniformly differentiable', and 'derivative' for 'uniform derivative'.

3.01. The relation between $f(x)$ and $f'(x)$ may be written in the form

$$f(X)-f(x) = (X-x)f'(x)+(X-x)0(r) \tag{D*}$$

obtained from equation (D) by multiplying both sides by $X-x$. Since (D) holds for any X near x but *different* from x, this multiplication by $X-x$ is permissible, but since both sides of the equation (D*) vanish when $X = x$, (D*) *is true even when* $X = x$.

3.02. A uniform derivative is a continuous function.

Since $f'(x) = \frac{f(X)-f(x)}{X-x} + 0(r)$ provided only $X-x = 0(s)$, we may interchange x and X in this equation, for this interchange does not alter the condition $X-x = 0(s)$ and therefore

$$\begin{aligned} f'(X) &= \frac{f(x)-f(X)}{x-X} + 0(r) \\ &= \frac{f(X)-f(x)}{X-x} + 0(r), \end{aligned}$$

so that $f'(X)-f'(x) = 0(r)-0(r) = 0(r-1)$, which proves that $f'(x)$ is continuous.

3.03. If $f(x)$ is differentiable in (a, b) then $f(x)$ is continuous in (a, b).

By 3.02, $f'(x)$, the derivative of $f(x)$, is continuous in (a, b) and so, by 2.1, $f'(x)$ is bounded and we can find a number m so that

$|f'(x)| < 10^m$. Hence since $f(X)-f(x) = (X-x)f'(x)+(X-x)0(1)$ provided $X-x = 0(n(1))$ it follows that

$$|f(X)-f(x)| \leqslant |X-x||f'(x)+0(1)| < |X-x| \,.\, 10^{m+1}$$

and so $f(X)-f(x) = 0(r)$ provided $X-x = 0(r+m+1)$, which proves that $f(x)$ is continuous.

3.04. If $f(x)$ is differentiable and if $x_n \to x$ then $\dfrac{f(x_n)-f(x)}{x_n-x} \to f'(x)$.

For we can choose ν so that $x_n-x = 0(n(r))$ provided $n \geqslant \nu$ and therefore $\dfrac{f(x_n)-f(x)}{x_n-x} = f'(x)+0(r)$ provided $n \geqslant \nu$, which proves that $\dfrac{f(x_n)-f(x)}{x_n-x} \to f'(x)$.

3.05. If $f(x)$ is differentiable and if $X_n \to x$ and $x_n \to x$ then

$$\frac{f(X_n)-f(x_n)}{X_n-x_n} \to f'(x).$$

Since $X_n \to x$ and $x_n \to x$, therefore $X_n-x_n \to 0$ and so we can choose ν such that $X_n-x_n = 0(n(r))$ if $n \geqslant \nu$. Therefore

$$\frac{f(X_n)-f(x_n)}{X_n-x_n} = f'(x_n)+0(r)$$

provided $n \geqslant \nu$. But $f'(x)$ is continuous so that $f'(x_n) \to f'(x)$, i.e. $f'(x_n) = f'(x)+0(r)$ for $n \geqslant$ some μ. Hence if N is the greater of μ and ν, $\dfrac{f(X_n)-f(x_n)}{X_n-x_n} = f'(x)+0(r-1)$ for $n \geqslant N$ and so

$$\frac{f(X_n)-f(x_n)}{X_n-x_n} \to f'(x).$$

3.1. The derivative of the sum of two functions is the sum of the derivatives of the functions.

Let $f'(x)$, $g'(x)$ be the derivatives of $f(x)$ and $g(x)$. Then

$$\frac{\{f(X)+g(X)\}-\{f(x)+g(x)\}}{X-x} = \frac{f(X)-f(x)}{X-x}+\frac{g(X)-g(x)}{X-x}$$
$$= f'(x)+g'(x)+0(r)+0(r) = f'(x)+g'(x)+0(r-1)$$

provided X is close enough to x, proving that $f'(x)+g'(x)$ is the derivative of $f(x)+g(x)$.

3.11. Theorem 3.1 readily extends to any number of functions. For instance, if $f'(x)$, $g'(x)$, $h'(x)$ are the derivatives of $f(x)$, $g(x)$,

$h(x)$ then by 3.1, $\{f'(x)+g'(x)\}$ is the derivative of $\{f(x)+g(x)\}$ and so, by applying 3.1 to the two functions $\{f(x)+g(x)\}$ and $h(x)$ it follows that $\{f'(x)+g'(x)\}+h'(x)$, i.e. $f'(x)+g'(x)+h'(x)$, is the derivative of $\{f(x)+g(x)\}+h(x)$.

3.2. If $f'(x)$, $g'(x)$ are the derivatives of $f(x)$, $g(x)$ then the derivative of the product $f(x)g(x)$ is $f'(x)g(x)+f(x)g'(x)$.

Since $f(x)$, $g(x)$ are differentiable, they are continuous and so bounded, and $g'(x)$ is also continuous and bounded, and therefore $f(X)+g(x)+g'(x)$ is bounded, for any x, X.

Now

$$\begin{aligned} f(X)g(X)-f(x)g(x) &= f(X)\{g(X)-g(x)\}+\{f(X)-f(x)\}g(x) \\ &= (X-x)[f(X)g'(x)+f'(x)g(x)+\{f(X)+g(x)\}0(r)] \end{aligned}$$

provided X is sufficiently close to x.

But, for such values of X, $f(X) = f(x)+0(r)$ and so

$$\frac{f(X)g(X)-f(x)g(x)}{X-x}-\{f(x)g'(x)+f'(x)g(x)\} = (g'(x)+f(X)+g(x))0(r)$$

which proves that $f(x)g'(x)+f'(x)g(x)$ is the derivative of $f(x)g(x)$ (by 1.31).

3.21. From 3.2 it follows that the derivative of $f(x)g(x)h(x)$ is

$$f'(x)g(x)h(x)+f(x)g'(x)h(x)+f(x)g(x)h'(x),$$

for

$$\begin{aligned} (f(x)g(x)h(x))' &= (\{f(x)g(x)\}h(x))' \\ &= (f(x)g(x))'h(x)+(f(x)g(x))h'(x) \\ &= (f'(x)g(x)+f(x)g'(x))h(x)+(f(x)g(x))h'(x) \\ &= f'(x)g(x)h(x)+f(x)g'(x)h(x)+f(x)g(x)h'(x), \end{aligned}$$

and step by step the theorem extends to products of any number of functions.

3.3. If, for any x (in some interval), $|f(x)| \geqslant 1/10^k$, for some constant k, and if $f'(x)$ is the derivative of $f(x)$ then the derivative of $1/f(x)$ is $-f'(x)/\{f(x)\}^2$.

Since $f(x)$ and $f'(x)$ are continuous they are bounded and so for some p, $|f'(x)|+|f(x)|$ is less than 10^p. Furthermore, if X is sufficiently close to x, $f(X)-f(x)$ and $\dfrac{f(X)-f(x)}{X-x}-f'(x)$ are both $0(r)$.

Now

$$\begin{aligned}
&\{1/f(X)-1/f(x)\}/(X-x)-(-f'(x)/f(x)^2)\\
&\quad=-\{(f(X)-f(x))/f(X)f(x)(X-x)\}+f'(x)/f(x)^2\\
&\quad=-\{f'(x)+0(r)\}/f(x)f(X)+f'(x)/f(x)^2\\
&\quad=-[f'(x)\{f(x)-f(X)\}+f(x)0(r)]/f(X)f(x)^2\\
&\quad=[\{f'(x)+f(x)\}/f(X)f(x)^2]0(r)
\end{aligned}$$

which proves the theorem since $\{f'(x)+f(x)\}/f(X)f(x)^2$ is bounded.

3.31. It follows from 3.2 and 3.3 that provided $|g(x)| \geqslant 1/10^k$ the derivative of $f(x)/g(x)$ is $\{f'(x)g(x)-f(x)g'(x)\}/g(x)^2$, for

$$\begin{aligned}
(f(x)/g(x))' &= f'(x)(1/g(x))+f(x)(1/g(x))'\\
&= f'(x)/g(x)-f(x)g'(x)/g(x)^2\\
&= \{f'(x)g(x)-f(x)g'(x)\}/g(x)^2.
\end{aligned}$$

Illustrations. If $f(x) = x$ then $\dfrac{f(X)-f(x)}{X-x} = \dfrac{X-x}{X-x} = 1$ and so the derivative of the function x is unity. This of course says no more than that 'x increases as fast as x'.

If for any x, or any x in some interval (a, b), we have $f(x) = f(a)$ then $f'(x) = 0$, for $\dfrac{f(X)-f(x)}{X-x} = \dfrac{f(a)-f(a)}{X-x} = 0$ for any X and x.

This result is usually expressed by saying that the derivative of a *constant* is zero, a constant being a function which has a value independent of the argument, i.e. one for which $f(X) = f(x)$ whatever values X, x may have. Such a function is for instance a number like 3, or to be more precise the function $f(x)$ defined to have the value 3 for any x, or any x in some interval.

It follows from 3.2 that the derivative of $f(x)g(x)$, when $f(x)$ is a constant, is just $f(x)g'(x)$; if c is the constant value of $f(x)$ then we may express this by saying that the derivative of $cg(x)$, for any number c whatsoever, is $cg'(x)$.

Thus the derivative of ax^2+bx+c is $2ax+b$, whatever numbers a, b, c may be, that is, whatever constants a, b, c may be, for the derivatives of ax^2, bx, c are $2ax$, b, 0 respectively and the derivative of the sum is the sum of the derivatives.

The derivative of x^n, where n is a positive whole number, is nx^{n-1}, for the derivative of x^1 is $1.x^0 = 1$, that of x^2 is $2x^1 = 2x$, and if for some n the derivative of x^n is nx^{n-1} then since $x^{n+1} = x.x^n$

the derivative of x^{n+1} is $1.x^n+x.nx^{n-1} = (n+1)x^n$, which proves the result for any positive whole number. When $n = -m$, $m > 0$, we have $x^n = 1/x^m$ and since the derivative of $1/x^m$ is

$$-mx^{m-1}/(x^m)^2 = -mx^{-m-1}$$

we see that the derivative of x^n is nx^{n-1} whether n is positive or negative.

The results we have just obtained may be summarized by saying that the derivative of

$$a_0+a_1x+a_2x^2+...+a_nx^n+b_1/x+b_2/x^2+...+b_m/x^m$$

is

$$a_1+2a_2x+3a_3x^2+...+na_nx^{n-1}-$$
$$-(b_1/x^2+2b_2/x^3+3b_3/x^4+...+mb_m/x^{m+1}),$$

in any interval not including the origin.

3.32. The derivative of $f(g(x))$ is $f'(g(x))g'(x)$.

Since $f'(x)$ and $g'(x)$ are continuous $f'(g(x))+g'(x)$ is bounded.

Now $$g(X) = g(x)+(X-x)(g'(x)+0(r))$$
and $$f(Y) = f(y)+(Y-y)(f'(y)+0(r))$$
provided $X-x$ and $Y-y$ are small enough.

Take $g(x)$ for y and $g(X)$ for Y so that $Y-y = (X-x)\{g'(x)+0(r)\}$, which can be made as small as we please by taking X close enough to x, and so

$$f(g(X))-f(g(x)) = (X-x)\{g'(x)+0(r)\}\{f'(g(x))+0(r)\}.$$

Therefore

$$\frac{f(g(X))-f(g(x))}{X-x}-f'(g(x))g'(x) = \{f'(g(x))+g'(x)+0(r)\}0(r),$$

but $f'(g(x))+g'(x)+0(r)$ is bounded, and so $f'(g(x))g'(x)$ is the derivative of $f(g(x))$.

Examples. To find the derivative of $\{x^2+3x+5\}^7$. Take

$$g(x) = x^2+3x+5 \quad \text{and} \quad f(x) = x^7$$

so that $(x^2+3x+5)^7 = f(g(x))$.

Since $f'(x) = 7x^6$ and $g'(x) = 2x+3$ therefore the derivative of $\{x^2+3x+5\}^7$ is $7(x^2+3x+5)^6(2x+3)$.

The derivative of $(x^3+4)^6/(x^2+2x+3)^5$ is

$$6(x^3+4)^5 3x^2/(x^2+2x+3)^5+(x^3+4)^6\{-5(x^2+2x+3)^{-6}(2x+2)\}$$
$$= 3(x^3+4)^5\{6x^2(x^2+2x+3)-5(x^3+4)(x+1)\}/(x^2+2x+3)^6$$
$$= 2(x^3+4)^5\{4x^4+13x^3+27x^2-20x-20\}/(x^2+2x+3)^6.$$

3.4. Implicit functions

We have already seen that the derivative of a constant function is zero and we shall now consider one of the many important parts which this result plays. Suppose for instance that we know that two functions $f(x)$ and $g(x)$, though defined in different ways, are equal for every value of x, can we say that these functions have equal derivatives? In other words, if we differentiate both sides of the identity $f(x) = g(x)$ do we still obtain an identity? Consider $f(x)-g(x)$; this difference is equal to zero for every value of x and is therefore a constant function and its derivative is zero. But the derivative of a difference of two functions is the difference of the derivatives and so $f'(x)-g'(x) = 0$ which proves that $f'(x) = g'(x)$.

We can now show that the derivative of $x^{p/q}$ *is* $(p/q)x^{p/q-1}$, i.e. *that the derivative of* x^n *is* nx^{n-1} *also when* n *is a fraction.*

For let $f(x) = x^{p/q}$ so that $\{f(x)\}^q = x^p$. *If we assume* that $f(x)$ has a derivative $f'(x)$, then the derivative of $\{f(x)\}^q$ is

$$q\{f(x)\}^{q-1}f'(x) = qx^{(p/q)(q-1)}f'(x) = qx^{p-p/q}f'(x),$$

and the derivative of x^p is px^{p-1} so that $qx^{p-p/q}f'(x) = px^{p-1}$, whence it follows that $f'(x) = (p/q)x^{p/q-1}$ as stated. The existence of the derivative of $x^{p/q}$ is proved in § 3.85. If instead of writing $f(x)$ for $x^{p/q}$ we write, for brevity, y, the relation between y and x is

$$y^q-x^p = 0,$$

and so y is a root of an equation in which one of the coefficients is a power of x. This is a particular case of the equation

$$a_0(x)y^m+a_1(x)y^{m-1}+a_2(x)y^{m-2}+...+a_{m-1}(x)y+a_m(x) = 0 \quad \text{(i)}$$

in which the coefficients $a_0(x)$, $a_1(x)$,..., $a_m(x)$ are polynomials in x. If we could solve this equation we should find m functions of x each of which when substituted for y would make the equation true for all values of x. For example, the equation

$$xy^2-(2x^2+x+2)y+(x^3+x^2+2x+2) = 0$$

has the roots $y = x+1$ and $y = (x^2+2)/x$ and each of these values of y makes the equation true for all values of x. It may happen that we are unable to solve the equation and in this case we say that the equation gives an *implicit* definition of m functions for which the equation is true; in such a case, of course, we know no more about these functions than we have stated, namely that they

satisfy a certain equation, but this alone often suffices to determine many other important properties of the functions and we shall see later that it suffices to determine the functions as power series. For the present we are concerned only to show that we can find the derivatives of functions defined implicitly, in terms of x and the functions themselves.

Let us think of y as a function of x for which the equation (i) is true for all values of x, so that the left-hand side of (i) is a constant function whose derivative is accordingly zero. Thus denoting by y' the derivative of the function for which y stands, assuming for the present that this function is derivable (an assumption which we shall justify in § 15.85), we have

$$a_0'(x)y^m+a_0(x)my^{m-1}y'+\ldots+a_{m-1}'(x)y+a_{m-1}(x)y'+a_m'(x) = 0$$

and therefore

$$y' = -\frac{\{a_0'(x)y^m+a_1'(x)y^{m-1}+\ldots+a_{m-1}'(x)y+a_m'(x)\}}{\{ma_0(x)y^{m-1}+(m-1)a_1(x)y^{m-2}+\ldots+2a_{m-2}(x)y+a_{m-1}(x)\}}. \tag{ii}$$

For instance, if

$$y^5-3xy+x^5 = 0 \quad \text{then} \quad 5y^4y'-3y-3xy'+5x^4 = 0$$

and so $$y' = (3y-5x^4)/(5y^4-3x).$$

Observe that although equation (ii) gives a unique value of the derivative y' this value itself depends upon y so that in general y' will have m values, one for each of the m values of y which satisfy equation (i). For example, if $y^2-(x^2+x+1)y+x^2(x+1) = 0$ we find $y' = \dfrac{3x^2+2x-(2x+1)y}{x^2+x+1-2y}$, but the equation is satisfied by $y = x+1$ and $y = x^2$, and for these values of y we find

$$y' = \frac{(3x^2+2x)-(2x+1)(x+1)}{x^2+x+1-2(x+1)} = 1$$

and $$y' = \frac{3x^2+2x-x^2(2x+1)}{x^2+x+1-2x^2} = 2x,$$

as we should expect.

EXAMPLES. The derivative of $x/\sqrt{(a^2+x^2)}$ is

$$1.(a^2+x^2)^{-\frac{1}{2}}+x\{-\tfrac{1}{2}(a^2+x^2)^{-\frac{3}{2}}2x\}$$
$$= (a^2+x^2)^{-\frac{3}{2}}\{a^2+x^2-x^2\} = a^2/(a^2+x^2)^{\frac{3}{2}}.$$

Alternatively write $y = x/\sqrt{(a^2+x^2)}$, then $(a^2+x^2)y^2 = x^2$ and so

$$2xy^2+2y(a^2+x^2)y' = 2x,$$

whence $$y(a^2+x^2)y' = x(1-y^2) = a^2x/(a^2+x^2)$$

and so $$y' = a^2/(a^2+x^2)^{\frac{3}{2}}.$$

To find the derivative of $\dfrac{\sqrt{(1+x)}-\sqrt[3]{(1-x)}}{\sqrt{(1+x)}+\sqrt[3]{(1-x)}}$ write y for the expression and we find

$$\frac{1+y}{1-y} = \frac{\sqrt{(1+x)}}{\sqrt[3]{(1-x)}} = \frac{(1+x)^{\frac{1}{2}}}{(1-x)^{\frac{1}{3}}} \quad \text{and so} \quad \left(\frac{1+y}{1-y}\right)^6 = \frac{(1+x)^3}{(1-x)^2},$$

i.e. $$(1+y)^6(1-x)^2 = (1-y)^6(1+x)^3,$$

whence

$$6(1+y)^5(1-x)^2y'-2(1+y)^6(1-x)$$
$$= -6(1-y)^5y'(1+x)^3+3(1-y)^6(1+x)^2$$

and therefore

$$6y'\left\{\left(\frac{1+y}{1-y}\right)^5(1-x)^2+(1+x)^3\right\} = (1-y)\left\{2\left(\frac{1+y}{1-y}\right)^6(1-x)+3(1+x)^2\right\},$$

i.e.

$$6y'\{(1-x)^{\frac{1}{3}}+(1+x)^{\frac{1}{2}}\}(1+x)^{\frac{5}{2}}$$
$$= (1-y)\{2(1+x)+3(1-x)\}(1+x)^2/(1-x),$$

i.e. $$y' = (5-x)/3(1-x)^{\frac{2}{3}}(1+x)^{\frac{1}{2}}\{(1-x)^{\frac{1}{3}}+(1+x)^{\frac{1}{2}}\}^2.$$

3.5. If the derivative of a function is positive the function is increasing, i.e. if $f'(x) > 0$ then $f(X) > f(x)$ when $X > x$ for any X near enough to x.

Since $f'(x) > 0$ we can choose r so that $f'(x) > 1/10^r$ and so $\dfrac{f(X)-f(x)}{X-x} = f'(x)+0(r) > 0$, which shows that $f(X) > f(x)$ when $X > x$ and $f(x) > f(X)$ when $x > X$.

Similarly, if the derivative is negative the function is decreasing, for we can choose r so that $f'(x) < -1/10^r$, making $f'(x)+0(r) < 0$ and so $f(X) < f(x)$ when $X > x$, and $f(x) < f(X)$ when $x > X$.

3.6. The expression $\dfrac{f(b)-f(a)}{b-a}$ is called the *mean slope* of $f(x)$ in (a, b) and will be denoted for short by $\mu(a, b)$ or $\mu(b, a)$.

If an interval (a, b) is divided into any number of equal parts

then in at least one of these parts the mean slope of $f(x)$ is at least as great as the mean slope in (a, b), and in at least one part the mean slope is no greater than in (a, b).

For if $x_0 = a, x_1, x_2, \ldots, x_n = b$ are the points of subdivision so that

$$x_r - x_{r-1} = (b-a)/n \quad \text{for any } r$$

then

$$f(b)-f(a) = \{f(x_1)-f(x_0)\}+\{f(x_2)-f(x_1)\}+\ldots+\{f(x_n)-f(x_{n-1})\}$$

and so

$$\frac{n\{f(b)-f(a)\}}{b-a} = \frac{f(x_1)-f(x_0)}{x_1-x_0}+\frac{f(x_2)-f(x_1)}{x_2-x_1}+\ldots+\frac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}$$

$$\sum_1^n \mu(x_r, x_{r-1}).$$

Let $\mu(x_k, x_{k-1})$ be the least of the n numbers $\mu(x_r, x_{r-1})$, $r = 1, 2, \ldots, n$ (choosing the least such value of k if there is more than one), then $\sum_1^n \mu(x_r, x_{r-1}) \geqslant n\mu(x_k, x_{k-1})$ and therefore $\frac{n\{f(b)-f(a)\}}{b-a} \geqslant n\mu(x_k, x_{k-1})$, i.e. $\mu(a, b) \geqslant \mu(x_{k-1}, x_k)$. Similarly, if $\mu(x_{g-1}, x_g)$ is the greatest of the n numbers

$$\mu(x_{r-1}, x_r), \quad r = 1, 2, \ldots, n,$$

then

$$\mu(a, b) \leqslant \mu(x_{g-1}, x_g).$$

3.61. $f'(x)$ is the derivative of $f(x)$ in (a, b). Then there are points c_1, c_2 in (a, b) such that $f'(c_1) \leqslant \mu(a, b)$ and $f'(c_2) \geqslant \mu(a, b)$.

Divide (a, b) into two equal parts; by 3.6, in one of these parts, (a_1, b_1) say, $\mu(a, b) \geqslant \mu(a_1, b_1)$. Divide (a_1, b_1) into two equal parts and choose that part (a_2, b_2) where $\mu(a_1, b_1) \geqslant \mu(a_2, b_2)$; divide (a_2, b_2) into two parts and choose (a_3, b_3), and so on. Since $\mu(a_n, b_n) \geqslant \mu(a_{n+1}, b_{n+1})$ for any n, therefore $\mu(a, b) \geqslant \mu(a_n, b_n)$ for any n. By Theorem 1.96 (a_n) and (b_n) tend to a common limit c_1 (say), and therefore by Theorem 3.05, $\mu(a_n, b_n) \to f'(c_1)$ and so $\mu(a, b) \geqslant f'(c_1)$. Similarly there is a point c_2 where $f'(c_2) \geqslant \mu(a, b)$.

3.62. If $f'(x) > \alpha \geqslant 0$ in (a, b) then $\mu(a, b) > \alpha$, for by 3.61 there is a point c where $\mu(a, b) \geqslant f'(c)$ and $f'(c) > \alpha$. If x, X are any two points in (a, b) the conditions of the theorem necessarily hold also in the interval (x, X) and therefore $\mu(x, X) > \alpha$.

3.621. In particular from $f'(x) > 0$ in (a, b) we derive $\mu(x, X) > 0$, i.e. $\dfrac{f(X)-f(x)}{X-x} > 0$, which shows that if $X > x$ then $f(X) > f(x)$ and so $f(x)$ is steadily increasing in (a, b). Similarly, if $f'(x) \geqslant 0$ in (a, b) then $f(X) \geqslant f(x)$ when $X \geqslant x$.

3.63. The closed interval (a, b) in 3.621 may be replaced by an open interval $[a, b]$.

3.64. If $f(x)$ is continuous in (a, b) and steadily increasing in the open interval $[a, b]$ then $f(x)$ is steadily increasing in (a, b).

We require to show that if x lies in $[a, b]$ then $f(a) < f(x) < f(b)$.

Since $f(x)$ is increasing in $[a, b]$ therefore $f\left(a+\dfrac{1}{n+1}\right) < f\left(a+\dfrac{1}{n}\right)$; but, as $f(x)$ is continuous in (a, b), $f\left(a+\dfrac{1}{n}\right) \to f(a)$ and therefore, by 1.47, $f(a) < f\left(a+\dfrac{1}{n}\right)$. If x lies in (a, b) we can choose n so that $a+\dfrac{1}{n} < x$ and therefore $f(a) < f\left(a+\dfrac{1}{n}\right) < f(x)$. Similarly

$$f(x) < f(b).$$

3.641. Theorems 3.63 and 3.64 together prove:

If $f(x)$ is continuous in (a, b) and $f'(x) > 0$ in $[a, b]$ then $f(x)$ is steadily increasing in (a, b).

3.65. The only function whose derivative is zero throughout an interval is a constant, i.e. if $f'(x) = 0$ for any x in (a, b) then $f(x) = f(a)$ for any x in (a, b).

Since $f(x)$ is differentiable

$$f(X)-f(x) = (X-x)\{f'(x)+0(r)\} \quad \text{provided } X-x = 0(s),$$

and so if $f'(x) = 0$,

$$f(X)-f(x) = (X-x)0(r).$$

Choose k so that $b-a < 10^k$ and let $N = 10^{k+s}$. Divide (a, b) into N equal parts by the points $a = a_0, a_1, a_2, \ldots, a_N = b$ so that the length of each interval (a_p, a_{p+1}) is $(b-a)/N = 0(s)$ and therefore

$$f(a_{p+1})-f(a_p) = (a_{p+1}-a_p)0(r) = \frac{b-a}{N}0(r).$$

Now any x between a and b necessarily falls into one of the N parts; suppose it falls between a_p and a_{p+1} so that $a_p < x \leqslant a_{p+1}$.

Since $x-a_p \leqslant a_{p+1}-a_p = 0(s)$, therefore

$$f(x)-f(a_p) = (x-a_p)0(r) = \frac{b-a}{N}0(r)$$

as $x-a_p \leqslant (b-a)/N$. Hence

$$\begin{aligned} f(x)-f(a) &= \{f(a_1)-f(a_0)\}+\{f(a_2)-f(a_1)\}+\ldots+\{f(a_p)-f(a_{p-1})\}+ \\ &\qquad\qquad +\{f(x)-f(a_p)\} \\ &= (p+1)\frac{(b-a)}{N}0(r) = (b-a)0(r), \quad \text{since } p+1 \leqslant N \\ &= 0(r-k). \end{aligned}$$

Since k is fixed and r may be as great as we please it follows that $f(x)$ and $f(a)$ are equal, to as many decimal places as we please, and therefore $f(x) = f(a)$ and $f(x)$ is constant.

3.651. If the derivatives of $f(x)$ and $g(x)$ are equal then $f(x)$ and $g(x)$ are equal apart from an additive constant, i.e. a function is uniquely determined by its derivative, apart from a constant. For if $\phi(x) = f(x)-g(x)$ then $\phi'(x) = f'(x)-g'(x) = 0$ and so, by 3.65, $\phi(x)$ is constant and therefore $f(x) = g(x)+\text{constant}$.

3.66. The foregoing proof shows also that if $f'(x) = 0(n)$ instead of $f'(x) = 0$ as in 3.65 then $f(x)-f(a) = 0(n-k)+0(r-k)$ for any r, i.e. $f(x)-f(a)-0(n-k)$ is zero to as many decimal places as we please and therefore $f(x)-f(a) = 0(n-k)$.

In particular if $f(a) = 0$ and $f'(x) = 0(n)$ then $f(x) = 0(n-k)$.

3.661. Theorem 3.65 may also be deduced from Theorem 3.61. For if x is any point in (a, b) then we can find c_1, c_2 in (a, x) such that

$$f'(c_1) \leqslant \frac{f(a)-f(x)}{a-x} \leqslant f'(c_2);$$

but $f'(c_1) = f'(c_2) = 0$, so that $f(a) = f(x)$ for all x in (a, b).

Similarly, if $f'(x) = 0(n)$ throughout (a, b), then from

$$f'(c_1) \leqslant \frac{f(x)-f(a)}{x-a} \leqslant f'(c_2)$$

we derive $f(x)-f(a) = (x-a)0(n) = 0(n-k)$, since

$$(x-a) \leqslant (b-a) < 10^k.$$

3.7. Derivative of a power series

We have proved that the derivative of a sum of any finite number of functions is the sum of the derivatives of the functions, but this theorem does *not* extend to the sum of an unlimited number of functions, except under very restrictive conditions. The following is the simplest result on the derivative of the sum of an unlimited number of functions.

If $\sum a_n R^n$ is convergent, then $\sum na_n x^{n-1}$ is the derivative of $\sum a_n x^n$ in the interval $(-x^*, x^*)$, where x^* is any positive number less than $|R|$.

Proof. Write $s_n(x) = \sum_{p=0}^{n} a_p x^p$, $\sigma_n(x) = \sum_{p=1}^{n} pa_p x^{p-1}$; by Theorem 1.83, $\sigma_n(x)$ is convergent in the interval $(-x^*, x^*)$. Let $s(x)$ and $\sigma(x)$ be the limits of $s_n(x)$ and $\sigma_n(x)$ respectively, so that, by Theorem 1.92, we can determine n, depending upon r but not upon x, such that

$$s(x) = s_n(x)+0(r), \qquad \sigma(x) = \sigma_n(x)+0(r)$$

for all x in $(-x^*, x^*)$. Further let $g(x)$ be a function whose derivative is $\sigma(x)$. Since $pa_p x^{p-1}$ is the derivative of $a_p x^p$ it follows that $\sigma_n(x) = s'_n(x)$. Hence, if $h_n(x) = g(x)-s_n(x)$, then

$$h'_n(x) = g'(x)-s'_n(x) = \sigma(x)-\sigma_n(x) = 0(r),$$

whence by Theorem 3.66, $h_n(x)-h_n(0) = 0(r-k)$, where k is chosen so that $2x^* < 10^k$. But $s(x)-s_n(x) = 0(r)$ and $h_n(0) = g(0)-a_0$, and therefore

$$g(x)-s(x) = h_n(x)-(s(x)-s_n(x)) = g(0)-a_0+0(r-k-1);$$

since this holds for *any* r, it follows that $s(x) = g(x)-g(0)+a_0$, and therefore $s'(x) = g'(x) = \sigma(x)$, which completes the proof.

We may express Theorem 3.7 by saying that the derivative of the limit of $s_n(x)$ is equal to the limit of the derivative of $s_n(x)$.

Observe that the result holds only in the *open* interval $[-R, R]$ and does not extend to the point R itself (unless $\sum na_n x^{n-1}$ is known to be convergent, when $x = R$, independently of Theorem 1.83). For instance, the series $\sum x^n/n^2$ is convergent when $x = 1$, but the derived series $\sum x^{n-1}/n$ is convergent only when x is numerically less than unity, and is divergent when $x = 1$.

It follows, of course, that if $\sum a_n x^n$ is convergent for every value of x, then its derivative is $\sum na_{n-1} x^{n-1}$ for every value of x.

The foregoing proof is based on the assumption that there is a function $g(x)$ whose derivative is the given continuous function $\sigma(x)$. The justification of this assumption is given in Chapter XV. The following is an alternative proof of Theorem 3.7 which is independent of this assumption.

3.701. *Alternative proof of* 3.7. By Theorem 1.83 $\sum na_n x^{n-1}$ is absolutely convergent in $(-x^*, x^*)$ and so, by 1.921, we can find n_r, independent of x, such that

$$\sum_n^N pa_p x^{p-1} = 0(r)$$

for all $N \geqslant n \geqslant n_r$, and all x in $(-x^*, x^*)$.

Let x, X be two points in $(-x^*, x^*)$ and write $\phi(x) = \sum_n^N a_p x^p$, so that $\phi'(x) = \sum_n^N pa_p x^{p-1}$. It follows from Theorem 3.61 that we can find c_1 and c_2 between x and X such that

$$\phi'(c_1) \leqslant \frac{\phi(X)-\phi(x)}{X-x} \leqslant \phi'(c_2);$$

but $\phi'(c_1) = 0(r)$ and $\phi'(c_2) = 0(r)$, and therefore

$$\frac{\phi(X)-\phi(x)}{X-x} = 0(r),$$

i.e.
$$\frac{1}{X-x}\sum_n^N a_p(X^p-x^p) = 0(r)$$

for all N and n not less than n_r, and therefore, since $\sum a_p x^p$ is convergent in $(-x^*, x^*)$, we have

$$\frac{1}{X-x}\sum_{p\geqslant n} a_p(X^p-x^p) = 0(r) \quad (n \geqslant n_r).$$

Furthermore, since $\sum na_n x^{n-1}$ converges in $(-x^*, x^*)$ we can find $\nu_r \geqslant n_r$, such that $\sum_{p\geqslant n} pa_p x^{p-1} = 0(r)$, for $n \geqslant \nu_r$ and x in $(-x^*, x^*)$, and we can determine λ_r, independent of x, so that if $s_n(x) = \sum_0^{n-1} a_n x^n$, then $\{s_{\nu_r}(X)-s_{\nu_r}(x)\}/(X-x)-s'_{\nu_r}(x) = 0(r)$ provided $X-x = 0(\lambda_r)$

Hence, if $f(x) = \sum a_n x^n$, then

$$\frac{f(X)-f(x)}{X-x} - \sum na_n x^{n-1}$$

$$= \sum_0^{\nu_r-1} a_p\left\{\left(\frac{X^p-x^p}{X-x}\right)-px^{p-1}\right\}+\sum_{p\geqslant\nu_r} a_p\left(\frac{X^p-x^p}{X-x}\right)-\sum_{p\geqslant\nu_r} pa_p x^{p-1}$$

$$= 0(r)+0(r)+0(r), \quad \text{provided } X-x = 0(\lambda_r),$$

$$= 0(r-1)$$

which proves that $f(x)$ is differentiable with derivative $\sum na_n x^{n-1}$ in $(-x^*, x^*)$.

3.71. If $f(x)$ is the derivative of $F(x)$, in $[-R, R)$, and $f(x) = \sum a_n x^n$, the power series being convergent for $x = R$, then

$$F(x)-F(0) = \sum a_n \frac{x^{n+1}}{n+1} \quad (|x| < |R|).$$

For by 1.912, $\sum a_n \frac{x^{n+1}}{n+1}$ is absolutely convergent for $|x| < |R|$, and $D\frac{x^{n+1}}{n+1} = x^n$, and therefore, by 3.7, $\sum a_n \frac{x^{n+1}}{n+1}$ is differentiable, with derivative $f(x)$, when $|x| < |R|$. Thus the difference

$$F(x)-\sum a_n \frac{x^{n+1}}{n+1}$$

is differentiable, with derivative zero, so that, by 3.65,

$$F(x)-\sum a_n \frac{x^{n+1}}{n+1} = F(0) \quad (|x| < |R|).$$

3.711.* The equation $F(x)-F(0) = \sum a_n \frac{x^{n+1}}{n+1}$ holds in fact not only for $|x| < |R|$, but also when $x = R$. For, by 1.911, $\sum a_n \frac{R^{n+1}}{n+1}$ is convergent, whence, by 2.6, both $\phi(x) = \sum a_n \frac{x^{n+1}}{n+1}$ and $F(x)$ are continuous in $(0, R)$; but $\phi(x) = F(x)-F(0)$ for $0 < x < R$ and so, by 2.7, $\phi(R) = F(R)-F(0)$.

3.72. The binomial theorem

If $\binom{m}{r}$ denotes $m(m-1)(m-2)...(m-r+1)/r!$ for any integer r, and any m, then $\left|\binom{m}{r}\Big/\binom{m}{r+1}\right| = \left|\frac{r+1}{r-m}\right| \to 1$, for fixed m.

Hence $\sum \binom{m}{r} x^r$ is absolutely convergent for all m, and all x in $[-1, 1]$.

Let $f(x) = \sum \binom{m}{r} x^r$ in $[-1, 1]$; then, by 3.7,

$$f'(x) = \sum \binom{m}{r} r x^{r-1} = m \sum \binom{m-1}{r} x^r$$

in $[-1, 1]$, and therefore

$$mf(x)-(1+x)f'(x) = \sum m\binom{m}{r}x^r - \sum m\binom{m-1}{r}x^r - \sum m\binom{m-1}{r-1}x^r$$

$$= m\sum\left\{\binom{m}{r}-\binom{m-1}{r}-\binom{m-1}{r-1}\right\}x^r.$$

But

$$\binom{m-1}{r-1}+\binom{m-1}{r} = \frac{(m-1)(m-2)\dots(m-r+1)}{r!}\{r+(m-r)\} = \binom{m}{r},$$

and so $$mf(x)-(1+x)f'(x) = 0.$$

If $-1 < -\alpha \leqslant x$, we have, obviously, $1+x \geqslant 1-\alpha > 0$, and therefore in the interval $(-\alpha, \alpha)$, where $0 < \alpha < 1$, the function $g(x) = f(x)/(1+x)^m$ is differentiable, with derivative

$$g'(x) = \{(1+x)f'(x)-mf(x)\}/(1+x)^{m+1} = 0;$$

hence $g(x) = g(0) = 1$, and so $f(x) = (1+x)^m$ for $-\alpha \leqslant x \leqslant \alpha$.

Thus $$(1+x)^m = 1+\binom{m}{1}x+\binom{m}{2}x^2+\dots+\binom{m}{r}x^r+\dots$$

for any m, and any x in $[-1, 1]$.

The result is also true for $x = 1, m > -1$, and for $x = -1, m > 0$ (see Example 4.8).

3.8. Inverse functions

If a function $f(x)$ is regarded as transferring a point x to a point y, then a function $g(x)$ which transfers y back to x is called an *inverse* of $f(x)$. For instance, $x-2$ is an inverse of $x+2$, and $1/x$ is its own inverse.

Formally we define: If each of the functions $f(x)$, $g(x)$ has a unique value for each value of x, and if for all x

$$f(g(x)) = x, \tag{i}$$

then $f(x)$ is called the *inverse* function of $g(x)$.

If we write y for $g(x)$ the condition (i) is equivalent to the pair of equations $y = g(x)$, $x = f(y)$. This suggests that if $f(x)$ is the inverse of $g(x)$ then $g(x)$ is the inverse of $f(x)$, but this is not true without some further restriction on the functions. Equation (i) shows that $g(x)$ takes each of its values once only, for if

$$g(x_1) = g(x_2) = v,$$

say, then $f(v) = f(g(x_1)) = x_1$ and also $f(v) = f(g(x_2)) = x_2$, and so as $f(x)$ has a *unique* value for $x = v$, therefore $x_1 = x_2$. Hence if $g(x)$ is also the inverse of $f(x)$ it follows that $f(x)$ takes its values once only and this is more than we were given by the conditions which made $f(x)$ the inverse of $g(x)$.

If, however, we add the condition

$$f(x) \textit{ takes each of its values once only} \tag{i}$$

we can readily prove that $g(x)$ is the inverse of $f(x)$, for writing $f(x)$ in place of x in equation (i) we have

$$f\{g(f(x))\} = f(x).$$

Now $f(x)$ takes each value once only so that $f(X) = f(x)$ only if $X = x$ and therefore from $f\{g(f(x))\} = f(x)$ we conclude, for all x,

$$g(f(x)) = x, \tag{i'}$$

which proves that $g(x)$ is the inverse of $f(x)$.

Alternatively, instead of imposing an additional condition on $f(x)$, we can prove (i′) subject to the condition that

$$\textit{every number is a value of } g(x). \tag{ii}$$

For if $y = g(x)$ then $f(y) = f(g(x)) = x$ and so

$$g(f(y)) = g(x) = y$$

and this is true for *all* values of y as every number is a value of $g(x)$.

We could of course have anticipated the need for this condition on $g(x)$ as equation (i′) says that every number x is a value of $g(f)$, for the value $f(x)$ of the argument.

Observe that by condition (i) we prove (i′) and from (i′) we can deduce condition (ii), and therefore condition (ii) follows from condition (i). Equally, since condition (i) follows from (i′) and (i′) follows from the condition (ii), therefore (i) follows from (ii). Thus the conditions (i) and (ii) are equivalent, for each follows from the other.

If the equation (i) holds, not for all values of x but for all values in some interval (a, b), then $f(x)$ is said to be the inverse of $g(x)$ in (a, b). For instance $\sqrt{x}$ is the inverse of x^2 in $(0, b)$ for any $b > 0$, and vice versa.

3.81. If $f(x)$ is the inverse of $g(x)$ and *if* $f(x)$ and $g(x)$ are differentiable then $f'(g(x)) = 1/g'(x)$, for $f(g(x)) = x$ and therefore

$$f'(g(x))g'(x) = 1,$$

whence the result follows.

3.82. If $g(x)$ is continuous in (a, b) and steadily increasing, so that $g(X) > g(x)$ whenever $X > x$, then $g(x)$ has a unique continuous inverse in (a, b).

If Y is a number between $g(a)$ and $g(b)$ then, since $g(x)$ is continuous, we can, by 2.41, find an X between a and b such that $g(X) = Y$; as $g(x)$ is steadily increasing there can only be one number X such that $g(X) = Y$. Let $f(y)$ be the function which takes the value X when y takes the value Y, so that for any Y between $g(a)$ and $g(b)$ the value of $f(Y)$ is uniquely determined. Since $f(Y) = X$ and $Y = g(X)$, therefore

$$g(f(Y)) = Y$$

for *any* Y between $g(a)$ and $g(b)$, so that $g(x)$ is the inverse of $f(y)$ in the interval $(g(a), g(b))$ and therefore, as $g(x)$ takes each of its values once only, $f(y)$ is the inverse of $g(x)$ in the interval (a, b).

Of course $f(y)$ is also steadily increasing, for if $f(Y_1) = X_1$, $f(Y_2) = X_2$, and $Y_1 < Y_2$, since $f(y)$ takes its values once only, X_1 and X_2 are unequal, and therefore, as $g(x)$ is steadily increasing,

$$(X_1 - X_2)\{g(X_1) - g(X_2)\} > 0,$$

i.e.

$$(X_1 - X_2)(Y_1 - Y_2) > 0;$$

whence, since $Y_1 < Y_2$, it follows that $X_1 < X_2$ and $f(y)$ is steadily increasing. Moreover, every number between a and b is a value of $f(y)$, for if X is any number in (a, b) then $f(g(X)) = X$, and therefore, by Theorem 2.23, $f(y)$ is continuous in $(g(a), g(b))$.

3.83. If $g(x)$ is differentiable in (α, β) and if its derivative $g'(x)$ is greater than some positive 2λ for any x in (α, β) then $g(x)$ has a unique inverse $f(x)$ which is also differentiable.

Since $g(x)$ is differentiable it is continuous, and since $g'(x) > 2\lambda$, $g(x)$ is steadily increasing in (α, β) and so, by 3.82, $g(x)$ has a unique continuous inverse $f(x)$ in (α, β). Let $g(x) = y$, $g(X) = Y$ so that $x = f(y)$ and $X = f(Y)$. Since $f(y)$ is continuous, $X-x = 0(k)$ when $Y-y = 0(p_k)$. Hence we can take y, Y sufficiently close to ensure that $g(X)-g(x) = (X-x)\{g'(x)+0(r)\}$ and so we have

$$\frac{f(Y)-f(y)}{Y-y} = \frac{X-x}{g(X)-g(x)} = \frac{1}{g'(x)+0(r)},$$

for X, x are unequal when Y, y are unequal since $f(y)$ is increasing steadily.

Therefore

$$\frac{f(Y)-f(y)}{Y-y} - \frac{1}{g'(x)} = \frac{0(r)}{g'(x)\{g'(x)+0(r)\}} = \frac{1}{2\lambda^2}0(r)$$

for $g'(x)\{g'(x)+0(r)\} \geqslant 2\lambda(2\lambda+0(r)) \geqslant 2\lambda^2$, provided r is chosen so that $10^r\lambda > 1$. Since λ is fixed and r may be as great as we please this proves that $f(y)$ is differentiable and $1/g'(x)$ is its derivative, where $x = f(y)$.

3.831. In 3.82 we may replace the condition '$g(x)$ is steadily increasing' by '$g(x)$ is steadily decreasing' and in 3.83 we may replace the condition '$g'(x)$ is greater than some positive 2λ' by '$g'(x)$ is less than some negative 2μ', making the appropriate changes in the proofs.

3.84.* If $g(x)$ is continuous in (a, b) and if for any $\alpha > 0$,

$$g'(x) \geqslant k_\alpha > 0$$

in $(a+\alpha, b-\alpha)$, then $g(x)$ has a unique inverse in the *closed* interval (a, b).

Since $g'(x) > 0$ in $(a+\alpha, b-\alpha)$ for any $\alpha > 0$, therefore $g'(x) > 0$ in $[a, b]$.

By 3.63 and 3.64 $g(x)$ is steadily increasing in (a, b) and so by 3.82 $g(x)$ has a unique continuous inverse $f(y)$ in (a, b); by 3.83 $f(y)$ is differentiable and $1/g'(x)$ is its derivative, when

$$g(a+\alpha) \leqslant y \leqslant g(b-\alpha)$$

and α is as small as we please. It is *not* necessarily true that $f(y)$ is differentiable in the closed interval $(g(a), g(b))$. The condition $g'(x) \geqslant k_\alpha > 0$ may be replaced by $g'(x) \leqslant -k_\alpha < 0$.

3.841.* In particular if, in addition to the conditions of 3.84, $g(x)$ takes any value whatsoever when x lies in $[a, b]$, then $f(y)$ is defined, and its derivative is $1/g'(f(y))$ for *all* values of y. For $f'(y) = 1/g'(f(y))$ for any y in the interval $(g(x_1), g(x_2))$ for *any* x_1, x_2 in $[a, b]$ and therefore, since $g(x)$ takes any chosen value for an appropriate x in $[a, b]$, this equation is true in *every* interval and therefore true for *all* values of y.

3.85. If q is a positive integer then the function x^q has the derivative qx^{q-1}, which exceeds $q\alpha^{q-1}$ when x exceeds α; hence by 3.83 x^q has a unique *differentiable* inverse, which is denoted by $x^{1/q}$. Accordingly $x^{1/q}$ is defined for any integer q and any *positive* x. $x^{1/q}$ satisfies $(x^{1/q})^q = (x^q)^{1/q} = x$. The derivative of $x^{1/q}$ is

$$1/q(x^{1/q})^{q-1} = x^{1/q}/qx.$$

Defining $x^{p/q}$ to be $(x^p)^{1/q}$ it follows that the derivative of $x^{p/q}$ is

$$px^{p-1}(x^p)^{1/q}/qx^p = px^{p/q}/qx.$$

Since $a^{pq} = (a^p)^q = (a^q)^p$ for integral p, q and any a, therefore $(x^{1/q})^{pq} = \{(x^{1/q})^q\}^p = x^p$; but $(x^{1/q})^{pq} = \{(x^{1/q})^p\}^q$ and so

$$\{(x^{1/q})^p\}^q = x^p, \quad \text{whence} \quad (x^{1/q})^p = (x^p)^{1/q} = x^{p/q}.$$

Furthermore $(x^{1/pq})^{pq} = x$, and so $\{(x^{1/pq})^p\}^q = x$, wherefore

$$(x^{1/pq})^p = x^{1/q} \quad \text{and so} \quad x^{1/pq} = (x^{1/q})^{1/p}.$$

Similarly $x^{1/pq} = (x^{1/p})^{1/q}$. Hence

$$x^{pr/qr} = (x^{1/qr})^{pr} = \{(x^{1/qr})^r\}^p = [\{(x^{1/q})^{1/r}\}^r]^p = (x^{1/q})^p = x^{p/q},$$

which justifies the use of the fractional notation.

$x^{p/q}$ satisfies the index laws, for

$$(x^{p/q})^{r/s} = [\{(x^{1/q})^p\}^r]^{1/s} = [(x^{1/q})^{pr}]^{1/s} = [x^{pr/q}]^{1/s}$$
$$= [\{x^{pr}\}^{1/q}]^{1/s} = [x^{pr}]^{1/qs} = x^{pr/qs},$$

and

$$x^{p/q}.x^{r/s} = x^{ps/qs}.x^{qr/qs} = (x^{1/qs})^{ps}.(x^{1/qs})^{qr}$$
$$= (x^{1/qs})^{ps+qr} = x^{(ps+qr)/qs} = x^{p/q+r/s}.$$

Similarly

$$x^{p/q}/x^{r/s} = (x^{1/qs})^{ps}/(x^{1/qs})^{rq} = (x^{1/qs})^{ps-qr} = x^{p/q-r/s}.$$

Accordingly $x^{p/q}/x = x^{p/q-1}$ and the derivative of $x^{p/q}$ may be written in the form $(p/q)x^{p/q-1}$ as in 3.4.

3.9. We have so far tacitly supposed that the functions with which we have been concerned are functions of a single variable and that the derivative of a function records how much faster or slower than the change in its single argument is the change in the value of the function. If, however, a function contains more than one argument we may consider it as depending upon each of these arguments in turn and examine the ratio of its change in value to the change in any single one of the arguments. For instance the function ax^2 is a function both of a and x, but we may think of it as a function of x alone or of a alone and seek for its rate of change when only x changes or when only a changes. The important point is that here a and x are *independent* variables for which a change in one is quite independent of a change in the other. We might express this by saying that as far as changes in a are concerned x is a constant. Thus we separate entirely the two questions 'How much faster than x does ax^2 change?' and 'How much faster than a does ax^2 change?' To answer the first question we observe that ax^2 is a product of x^2 with a *constant* function a and so the rate of change which we require is given by the derivative $2ax$, but in the second question it is with changes in a that we are concerned, and so now ax^2 is a product of the function a with the *constant* x^2 and the rate of change required is now given by x^2, since the derivative of a is unity. The presence of more than one argument in the function makes a change necessary in the notation for a derivative. For instance, if we were to denote the derivative of a function of two arguments, $f(x, a)$, by $f'(x, a)$ we could not tell whether this derivative records the rate of change of $f(x, a)$ for changes in x or for changes in a, and as we have already seen, these rates of change might be quite different. Accordingly we introduce the notation $D_x f(x, a)$, $D_a f(x, a)$ to express the rate of change of $f(x, a)$ as x varies and as a varies, respectively. The extension of the notation to functions of three or more arguments is quite obvious, $D_x f(x, y, z)$, $D_y f(x, y, z)$, $D_z f(x, y, z)$ denoting the derivatives according as x, y, or z varies, and so on.

We proved in 3.65 that if $D_x f(x) = 0$, as x varies in some interval (a, b), then $f(x) = f(a)$ for any x in (a, b). It will readily be observed that nowhere in the proof did we assume that $f(x)$ is a function of a single variable and we may therefore express 3.65

in the following form: If $D_x f(x, y) = 0$ as x varies in (a, b) then $f(x, y) = f(a, y)$ for any x in (a, b), i.e. $f(x, y)$ is in fact a function of y alone, in the sense that the value of the function is independent of the value of x, in the interval (a, b).

If x and y are not independent variables but one of them, y, suppose, stands for some function of x, then we can no longer form the derivative of $f(x, y)$ by supposing that x alone varies and y is constant. For instance, if $y = x^2$, the value of $D_x yx^3$ is not $3yx^2$ but is $5x^4$. For the present we shall not have occasion to use a special notation to distinguish these cases but we shall return to this point in a later chapter.

There is a further notation for the derivative of a function which is of both historical and practical importance. This notation consists in writing d/dx in place of the D_x we introduced above, and d/dy in place of D_y, etc. In this notation many of the theorems we have proved on the derivatives of composite functions take a very simple form.

In 3.32 we proved that the derivative of $f(g(x))$ is $f'(g(x))g'(x)$. Let y denote $f(g(x))$ and t denote $g(x)$ so that $y = f(t)$. If we think of t as an independent variable we have $D_t y = f'(t)$, and since $t = g(x)$, $D_x t = g'(x)$, and therefore the fact that the derivative of y, i.e. of $f(g(x))$, is $f'(g(x))g'(x)$, is expressed by the equation

$$D_x y = D_t y . D_x t.$$

Now writing d/dx for D_x, etc., we have

$$\frac{dy}{dx} = \frac{dy}{dt} . \frac{dt}{dx}. \tag{i}$$

Thus equation (i) is the expression in this notation of Theorem 3.32. In 3.81 we showed that if $f(x)$ is the inverse of $g(x)$ then

$$f'(g(x)) = 1/g'(x);$$

write y for $g(x)$ so that $y = g(x)$ and $x = f(y)$. Then

$$f'(g(x)) = f'(y) = D_y x \quad \text{and} \quad g'(x) = D_x y$$

and so

$$D_y x = 1/D_x y,$$

i.e.

$$\frac{dx}{dy} = 1 \Big/ \frac{dy}{dx}. \tag{ii}$$

Equations (i) and (ii) show that the sign dy/dx *behaves* as if it were a quotient of the quantities dy, dx, for (i) is formed as if

the dt in the numerator and denominator on the right-hand side could be cancelled and (ii) says that dy/dx may be inverted like a fraction. Of course our definition shows that dy/dx is not a fraction, standing as it does for $D_x y$; in fact dy/dx is not composed of dy and dx but of d/dx and y, and d/dx is just a shorthand for the expression 'the rate of change as x varies'. Equations (i) and (ii) are useful brief expressions of Theorems 3.32 and 3.81, easy to remember and to apply, but nothing more. It must not be supposed that because dy/dx looks like a fraction, and because if it be so regarded equations (i) and (ii) are obviously true, that the proofs we have given of 3.32 and 3.81 are superfluous. On the contrary it is these very proofs which justify our saying that dy/dx behaves like a fraction, and nothing else. Incidentally this notation effects no simplification in the expression of Theorems 3.1 and 3.2. The origin of the notation lies in the relation of the derivative y' of a function of x to the ratio of 'difference in y' to 'difference in x', which might be written 'diff y/diff x' and the further abbreviation to dy/dx is then strongly suggested. The use of the term 'differentiating' as an alternative to 'deriving' is explained in the same way, for the word 'differentiating' is a synonym of 'differencing', meaning 'taking differences'.

IV

THE DERIVATION INVARIANT

THE EXPONENTIAL AND LOGARITHMIC FUNCTIONS. HYPERBOLIC FUNCTIONS

4. In this chapter we commence the application of the general results we have established to the study of a number of special functions. The first we consider is the *exponential* function $E(x)$ whose value for each value of x is the limit of the series

$$1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^n}{n!}+\ldots.$$

This series is absolutely convergent for every value of x, since

$$\frac{|x^{n+1}|}{(n+1)!}\Big/\frac{|x^n|}{n!}=\frac{|x|}{n}<\frac{1}{2}\quad\text{provided } n>2|x|.$$

Hence by 3.7 the derivative of $E(x)$ is $\sum\limits_{n=1}\frac{nx^{n-1}}{n!}$ for every x; but $\frac{nx^{n-1}}{n!}=\frac{x^{n-1}}{(n-1)!}$ and so *the derivative of $E(x)$ is the function $E(x)$ itself.* $E(x)$, being differentiable, is continuous.

Note that $E(0)=1$. The value $E(1)$ is denoted for brevity by e, so that e stands for the limit of the convergent series

$$1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\ldots$$

from which the value of e can be calculated to as many decimal places as we please. To 23 places

$$e=2{\cdot}718281\ 828459\ 045235\ 36028.$$

If $x\geqslant 0$, $E(x)\geqslant 1+x\geqslant 1$, and so $E(x)$ can be made as great as we please by choosing x sufficiently great.

4.1. We prove next the important property $\mathbf{E(x+a)=E(x).E(a)}$ for any x and a.

The equation is obviously true if $a=0$, since $E(0)=1$. We assume first that $x+a\geqslant 0$; then since $E(x+a)\geqslant 1$,

$$D_x\{1/E(x+a)\}=-E(x+a)/(E(x+a))^2=-1/E(x+a),$$

and therefore

$$D_x\{E(x).E(a)/E(x+a)\}$$
$$= [D_x E(x)E(a)]/E(x+a)+E(x)E(a)D_x(1/E(x+a))$$
$$= \{E(x)E(a)-E(x)E(a)\}/E(x+a) = 0,$$

for every x and a such that $x+a \geqslant 0$. Therefore by 3.65

$$E(x)E(a)/E(x+a) = E(0)E(a)/E(0+a) = 1, \quad \text{if } E(a) \neq 0,$$

but $E(a) \geqslant 1$ if $a \geqslant 0$, and so $E(x)E(a) = E(x+a)$ if $x+a \geqslant 0$ and $a \geqslant 0$; by symmetry the result therefore holds for $x+a \geqslant 0$ and $x \geqslant 0$, and so it holds provided only $x+a \geqslant 0$ since when $x+a \geqslant 0$, at least one of x and a is non-negative.

In particular $E(x)E(-x) = E(x-x) = E(0) = 1$ and so

$$\mathbf{E(-x) = 1/E(x)}$$

for any x. When x is positive $E(x)$ is positive and greater than unity and so $E(-x)$ lies between 0 and 1; in particular $E(x)$ is not zero for any value of x and $E(x) > 0$ for every x, positive or negative.

If $x+a < 0$ then $-x-a > 0$ and so

$$E(x+a) = 1/E(-x-a) = 1/E(-x).E(-a) = E(x)E(a),$$

which completes the proof that $E(x+a) = E(x).E(a)$ for all values of x and a.

4.11. $E(a+b+c) = E(a+b)E(c) = E(a)E(b)E(c)$ and so on for any number of factors.

4.12. $E(x)$ is steadily increasing for all values of x.

Choose any X_1, X_2 so that $X_2 > X_1$ and let $d = X_2-X_1$ so that $d > 0$. Then $E(X_2) = E(X_1+d) = E(d)E(X_1) > E(X_1)$ since $E(d) > 1$ and $E(X_1) > 0$ whatever X_1 may be, which proves that $E(x)$ is steadily increasing.

4.2. If n is a positive whole number $E(nx) = \{E(x)\}^n$.

For $nx = x+x+x+...+x$, where the sum has n terms, and so

$$E(nx) = E(x).E(x).E(x)...E(x), \quad \text{with } n \text{ factors in the product,}$$
$$= \{E(x)\}^n.$$

In particular if $x = 1$, $E(n) = \{E(1)\}^n = e^n$.

Furthermore

$$E(-nx) = 1/E(nx) = \{E(x)\}^{-n}$$

and so the equation $E(nx) = \{E(x)\}^n$ is true whether n be positive or negative.

4.21. For any fraction $\frac{p}{q}$, $E\left(\frac{p}{q}x\right) = \{E(x)\}^{p/q}$.

By 4.2
$$\left\{E\left(\frac{p}{q}x\right)\right\}^q = E\left(q\frac{p}{q}x\right) = E(px) = \{E(x)\}^p$$

and therefore
$$E\left(\frac{p}{q}x\right) = \{E(x)\}^{p/q}.$$

Thus the equation $E(nx) = \{E(x)\}^n$ is true whether n be a whole number or a fraction.

In particular $E(n) = e^n$ for all integral or fractional values of n, which shows why $E(x)$ is known as the *exponential* function.

4.3. The only function which is its own derivative is $aE(x)$.

For if $f(x)$ is a function which is its own derivative then

$$\begin{aligned} D_x(f(x)/E(x)) &= D_x(f(x)E(-x)) \\ &= (D_x f(x))E(-x)+f(x)(D_x E(-x)) \\ &= f(x)E(-x)-f(x)E(-x), \quad \text{since } D_x(-x) = -1, \\ &= 0. \end{aligned}$$

Therefore $f(x)/E(x) = f(0)/E(0) = f(0)$, i.e. $f(x) = f(0)E(x)$ and since $f(0)$ is just a constant we may denote it by a.

In particular if $f(x)$ is its own derivative and if $f(0) = 1$, then $f(x) = E(x)$.

4.4. We have seen that $E(x)$ is continuous and steadily increasing, and therefore, by 3.82, $E(x)$ has an inverse function; this inverse is called $\log x$ so that $\log(E(x)) = x$ and $E(\log x) = x$ and if

$$y = E(x) \quad \text{then} \quad x = \log y.$$

Since $E(x) > 0$ for every value of x the function $\log y$ is defined for *every positive value* of y and *only* for positive values of y. Moreover, also by 3.82, $\log y$ is *steadily* increasing.

4.41. The function $\log x$ is differentiable and its derivative is $1/x$.

For if N is any positive number, as large as we please, for any x in the interval $(-N, N)$ we have $E(x) \geqslant E(-N)$ since $E(x$ increases, and therefore $E(x) \geqslant 1/E(N) > 0$ and so

$$E'(x) \geqslant 1/E(N)$$

and Theorem 3.83 applies proving that $\log x$ is differentiable and that

$$D_y \log y = 1/E(\log y)$$

and so

$$D_y \log y = 1/y.$$

This is true in the interval $E(-N)$, $E(N)$ for any value of N; since $E(N)$ can be made as great as we please by choosing a great enough value of N, and since $E(-N) = 1/E(N)$ which accordingly can be made as small as we please, the interval $E(-N)$, $E(N)$ will contain any *positive* value of y we please, great or small, and therefore $1/y$ is the derivative of $\log y$ for *any positive value* of y.

4.5. Consider the expression $E(b \log a)$. When b is an integer or a fraction we see, by 4.21, that

$$E(b \log a) = \{E(\log a)\}^b.$$

But $E(\log a) = a$ and so $E(b \log a) = a^b$.

Although a^b is defined when b is an integer or a fraction there is no elementary definition of this function when b is an endless decimal. Since an endless decimal is the limit of a sequence of terminating decimals (for instance $\sqrt{2}$ is the limit of the sequence 1, 1·4, 1·41, 1·414,...) we define a^b, when b is the limit of a sequence of fractions $b_1, b_2, b_3, \ldots$, to be the limit of the sequence $a^{b_1}, a^{b_2}, a^{b_3}, \ldots$. To make this definition effective we must however prove that the sequence $a^{b_1}, a^{b_2}, a^{b_3}, \ldots$ is convergent.

Since each b_n is a fraction, $a^{b_n} = E(b_n \log a)$ and by means of this relation we readily prove the convergence of the sequence. For $E(x)$ is a continuous function and $b_n \log a \to b \log a$ and therefore $E(b_n \log a) \to E(b \log a)$, by Theorem 2.3. This proves both that a^{b_n} is convergent and also that its limit is $E(b \log a)$. This limit depends only on a and b, and is independent of the particular sequence $b_1, b_2, \ldots$ we used to find the limit. Since we defined a^b to be the limit of a^{b_n} we have proved that $a^b = E(b \log a)$ also when b is an endless decimal. It follows that

$$\log a^b = \log E(b \log a) = b \log a.$$

In particular if we give a the value e we see that

$$e^b = E(b \log e).$$

But $e = E(1)$ and so $\log e = \log E(1) = 1$, i.e. $\log e = 1$, and therefore

$$e^b = E(b).$$

Thus for any value of x, $E(x) = e^x$ and accordingly we shall write e^x instead of $E(x)$ from this point onwards.

4.51. Since $e^x > x^{p+1}/(p+1)!$ for any integral value of $p \geqslant 0$, and $x \geqslant 0$, therefore

$$x^p/e^x < (p+1)!/x < 1/10^r \quad \text{if } x > 10^r(p+1)!$$

In particular $$n^p/e^n \to 0.$$

Writing $\log n$ for x we have

$$(\log n)^p/n = 0(r) \quad \text{for } \log n > 10^r(p+1)!,$$

i.e. since e^x is steadily increasing, for $n > e^{10^r(p+1)!}$, which proves that $(\log n)^p/n \to 0$ for any $p \geqslant 0$.

From $(\log n)/n \to 0$ we deduce $(\log m^{p/q})/m^{p/q} \to 0$, by writing $m^{p/q}$ for n (where $p/q > 0$), whence it follows that $\{(p/q)\log m\}/m^{p/q} \to 0$ and therefore that $(\log m)/m^{p/q} \to 0$.

The theorem that $(\log n)/n \to 0$ may be written in the form

$$\log n^{1/n} \to 0,$$

from which it follows, since e^x is continuous, that

$$e^{\log n^{1/n}} \to e^0,$$

i.e. $$n^{1/n} \to 1.$$

4.6. Just as $\sum_0^n \frac{x^k}{k!}$ and its derivative $\sum_0^{n-1} \frac{x^k}{k!}$ are nearly equal (because $x^n/n! \doteq 0$) for large values of n, suggesting that the limit of $\sum_0^n \frac{x^k}{k!}$ is a function which is its own derivative, so too the function $\left(1+\frac{x}{n}\right)^n$ and its derivative $n\left(1+\frac{x}{n}\right)^{n-1}\frac{1}{n} = \left(1+\frac{x}{n}\right)^{n-1}$ are nearly equal for large values of n, for their ratio is $1+\frac{x}{n}$ which is nearly unity, and so since $\left(1+\frac{x}{n}\right)^n = 1$ when $x = 0$, we should expect that $\left(1+\frac{x}{n}\right)^n$ is convergent and that its limit is e^x, the only function which takes the value unity when $x = 0$ and which is its own derivative. We can easily show that this is the case, though the proof is somewhat indirect.

Let $y = 1+xt$ then

$$\frac{d}{dt}\log y = \frac{d}{dy}\log y\frac{dy}{dt} = \frac{1}{y}x = \frac{x}{1+xt}$$

and therefore

$$\frac{\log(1+xT)-\log(1+xt)}{T-t} = \frac{x}{1+xt}+0(r) \quad \text{provided } T-t = 0(s).$$

Take $t = 0$ and we have

$$\frac{\log(1+xT)}{T} = x+0(r) \quad \text{if } T = 0(s). \tag{i}$$

Let $T = 1/n$ and so

$$n\log\left(1+\frac{x}{n}\right) = x+0(r) \quad \text{provided } n > 10^{s},$$

i.e. $$n\log\left(1+\frac{x}{n}\right) \to x.$$

But e^x is continuous and therefore

$$e^{n\log(1+x/n)} \to e^x,$$

i.e. $$\left(1+\frac{x}{n}\right)^n \to e^x.$$

Similarly, if we take $T = -\frac{1}{n}$ in (i) we find $\left(1-\frac{x}{n}\right)^{-n} \to e^x$.

Thus both $\left(1+\frac{x}{n}\right)^n$ and $\left(1-\frac{x}{n}\right)^{-n}$ tend to e^x.

In particular if $x = 1$ we see that both

$$\left(1+\frac{1}{n}\right)^n \quad \text{and} \quad \left(1-\frac{1}{n}\right)^{-n} \quad \text{tend to } e.$$

This illustrates an important point in the theory of limits, for although $1+\frac{1}{n} \to 1$ and $\left(1+\frac{1}{n}\right)\left(1+\frac{1}{n}\right)\ldots$, to any fixed number of factors, also tends to 1 yet if the number of factors is not fixed this is no longer true, for as we have seen when there are n factors the limit is not unity but e, which is greater than 2.

An even simpler illustration of the same point is afforded by the sum $1/n+1/n+\ldots$; if there are a fixed number of terms in the sum, the limit of the sum is the sum of the limits of each term,

which is zero, but if there are n terms in the sum its value is unity, however great n may be.

EXAMPLES. (i) The derivative of x^n is nx^{n-1} for any positive value of x, whatever value n may have.

For $x^n = e^{n \log x}$ and so if $y = n \log x$

$$\frac{d}{dx}x^n = \frac{d}{dx}e^y = \frac{d}{dy}e^y.\frac{dy}{dx} = e^y.\frac{n}{x} = x^n.\frac{n}{x} = nx^{n-1}.$$

(ii) The derivative of $e^{f(x)}$ is $e^{f(x)}f'(x)$; for if $y = f(x)$,

$$\frac{d}{dx}e^{f(x)} \cdot \frac{d}{dy}e^y\frac{dy}{dx} = e^y f'(x) = e^{f(x)}f'(x).$$

(iii) The derivative of $\log f(x)$ is $f'(x)/f(x)$; for if $y = f(x)$,

$$\frac{d}{dx}\log f(x) = \frac{d}{dx}\log y = \frac{d}{dy}\log y.\frac{dy}{dx} = \frac{1}{y}f'(x) = \frac{f'(x)}{f(x)}.$$

(iv) To find the derivative of x^x. Let $y = x^x$ then $\log y = x \log x$. Now

$$\frac{d}{dx}\log y = \frac{d}{dy}\log y.\frac{dy}{dx} = \frac{1}{y}\frac{dy}{dx}$$

and

$$\frac{d}{dx}(x\log x) = \log x + x.\frac{1}{x} = 1+\log x,$$

and therefore

$$\frac{dy}{dx} = y(1+\log x) = x^x(1+\log x).$$

(v) We define $\log_a x = \log x/\log a$. Therefore

$$\log_a a^x = \log a^x/\log a = x\log a/\log a = x,$$

and so $\log_a x$ is the inverse function of a^x. Furthermore

$$a^{\log_a x} = e^{\log_a x.\log a} = e^{\log x} = x$$

and therefore a^x is the inverse function of $\log_a x$. Notice that since $\log e = 1$ therefore $\log_e x = \log x$.

It follows from the definition that

$$\frac{}{dx}\log_a x \cdot \frac{1}{x\log a}$$

and

$$\frac{d}{da}\log_a x = -\frac{\log x}{(\log a)^2}.\frac{1}{a} = -\frac{\log x}{a(\log a)^2}.$$

4.7. The hyperbolic functions

The functions $\frac{1}{2}(e^x+e^{-x})$ and $\frac{1}{2}(e^x-e^x)$ are known as the *hyperbolic cosine* and *hyperbolic sine* respectively, and are denoted by

ch x and sh x (pronounced 'c—h—x' and 's—h—x', not 'ch—x' and 'sh—x').†

Since $\frac{1}{2}(e^{-x}+e^{x}) = \frac{1}{2}(e^{x}+e^{-x})$ it follows that $\text{ch}(-x) = \text{ch}\,x$ and since $\frac{1}{2}(e^{-x}-e^{x}) = -\frac{1}{2}(e^{x}-e^{-x})$ therefore $\text{sh}(-x) = -\text{sh}\,x$.

Adding and subtracting the hyperbolic functions we obtain the very useful equations

$$\text{ch}\,x+\text{sh}\,x = e^{x}, \qquad \text{ch}\,x-\text{sh}\,x = e^{-x}.$$

Whether x be positive or negative, $(e^{x}-1)^2 \geqslant 0$, equality occurring only when $e^{x} = 1$, i.e. when $x = \log 1 = 0$, and therefore

$$e^{2x}+1 \geqslant 2e^{x};$$

since e^{x} is positive for any x we may divide by e^{x}, giving

$$e^{x}+e^{-x} \geqslant 2,$$

whence it follows that, for any x, $\text{ch}\,x \geqslant 1$, equality occurring only when $x = 0$. If x is positive $e^{x} > 1$ and so, since $e^{-x} = 1/e^{x}$, we have $e^{x} > e^{-x}$ and therefore $\text{sh}\,x > 0$.

If x is negative $\text{sh}\,x = -\text{sh}|x|$ and therefore $\text{sh}\,x < 0$.

If $\text{sh}\,x = 0$, $e^{x} = e^{-x}$ and so $e^{2x} = 1$, $2x = \log 1 = 0$, i.e. $x = 0$.

4.71. $\quad \text{ch}^2x-\text{sh}^2x = (\text{ch}\,x-\text{sh}\,x)(\text{ch}\,x+\text{sh}\,x) = e^{-x}.e^{x} = 1,$

i.e. for any value of x, $\text{ch}^2x-\text{sh}^2x = 1$, where ch^2x stands for $(\text{ch}\,x)^2$, etc.

4.711. Since $\text{ch}\,x > 0$ it follows from 4.71 that

$$\text{ch}\,x = +\sqrt{(1+\text{sh}^2x)};$$

the sign of sh x, however, depends upon the sign of x and therefore

$$\text{sh}\,x = +\sqrt{(\text{ch}^2x-1)} \quad \text{if } x \geqslant 0,$$
$$\text{sh}\,x = -\sqrt{(\text{ch}^2x-1)} \quad \text{if } x \leqslant 0.$$

4.72.
$$D_x\,\text{sh}\,x = D_x\,\tfrac{1}{2}(e^{x}-e^{-x}) = \tfrac{1}{2}(e^{x}+e^{-x}) = \text{ch}\,x,$$
$$D_x\,\text{ch}\,x = D_x\,\tfrac{1}{2}(e^{x}+e^{-x}) = \tfrac{1}{2}(e^{x}-e^{-x}) = \text{sh}\,x.$$

Thus each of sh x, ch x is the derivative of the other.

4.73. Addition formulae

$$\mathbf{sh(x+y) = sh\,x\,ch\,y+sh\,y\,ch\,x.}$$

† Many authors write sinh x, cosh x instead of sh x, ch x. In German works we find 'sin x', 'cos x' printed in Gothic characters.

For

$$4\,\mathrm{sh}\,x\,\mathrm{ch}\,y = (e^x - e^{-x})(e^y + e^{-y}) = e^{x+y} - e^{-x-y} + e^{x-y} - e^{-x+y}$$

and

$$4\,\mathrm{ch}\,x\,\mathrm{sh}\,y = (e^x + e^{-x})(e^y - e^{-y}) = e^{x+y} - e^{-x-y} - e^{x-y} + e^{-x+y}$$

and therefore, adding these results,

$$4(\mathrm{sh}\,x\,\mathrm{ch}\,y + \mathrm{ch}\,x\,\mathrm{sh}\,y) = 2(e^{x+y} - e^{-x-y}) = 4\,\mathrm{sh}(x+y),$$

whence the stated result follows.

From $\mathrm{sh}(x+y) = \mathrm{sh}\,x\,\mathrm{ch}\,y + \mathrm{ch}\,x\,\mathrm{sh}\,y$ we have

$$D_x\,\mathrm{sh}(x+y) = D_x[\mathrm{sh}\,x\,\mathrm{ch}\,y + \mathrm{ch}\,x\,\mathrm{sh}\,y],$$

i.e. $$\mathbf{ch(x+y) = ch\,x\,ch\,y + sh\,x\,sh\,y}.$$

Replacing y by $-y$ in these addition formulae we find, using

$$\mathrm{ch}(-y) = \mathrm{ch}\,y, \qquad \mathrm{sh}(-y) = -\mathrm{sh}\,y,$$

that

$$\mathrm{sh}(x-y) = \mathrm{sh}\,x\,\mathrm{ch}\,y - \mathrm{ch}\,x\,\mathrm{sh}\,y,$$
$$\mathrm{ch}(x-y) = \mathrm{ch}\,x\,\mathrm{ch}\,y - \mathrm{sh}\,x\,\mathrm{sh}\,y.$$

4.74. Replacing y by x in the addition formulae we obtain the *duplication formulae*

$$\begin{aligned}\mathrm{sh}\,2x &= 2\,\mathrm{sh}\,x\,\mathrm{ch}\,x,\\ \mathrm{ch}\,2x &= \mathrm{ch}^2x + \mathrm{sh}^2x\\ &= 2\,\mathrm{sh}^2x + 1, \quad \text{using 4.71},\\ &= 2\,\mathrm{ch}^2x - 1.\end{aligned}$$

4.75. The *hyperbolic tangent, cotangent, secant,* and *cosecant,* denoted by $\mathrm{th}\,x$, $\coth x$, $\mathrm{sech}\,x$, $\mathrm{cosech}\,x$ (pronounced t—h—x, cot—h—x, sec—h—x, cosec—h—x), are defined by the equations

$$\mathrm{th}\,x = \mathrm{sh}\,x/\mathrm{ch}\,x, \qquad \coth x = \mathrm{ch}\,x/\mathrm{sh}\,x,$$
$$\mathrm{sech}\,x = 1/\mathrm{ch}\,x, \qquad \mathrm{cosech}\,x = 1/\mathrm{sh}\,x.$$

From 4.71, dividing in turn by ch^2x and sh^2x we find,

$$\mathrm{sech}^2x = 1 - \mathrm{th}^2x \quad \text{and} \quad \mathrm{cosech}^2x = \coth^2x - 1.$$

From the addition formulae for $\mathrm{sh}\,x$ and $\mathrm{ch}\,x$ we obtain

$$\begin{aligned}\mathrm{th}(x+y) &= \mathrm{sh}(x+y)/\mathrm{ch}(x+y)\\ &= (\mathrm{sh}\,x\,\mathrm{ch}\,y + \mathrm{ch}\,x\,\mathrm{sh}\,y)/(\mathrm{ch}\,x\,\mathrm{ch}\,y + \mathrm{sh}\,x\,\mathrm{sh}\,y)\\ &= (\mathrm{th}\,x + \mathrm{th}\,y)/(1 + \mathrm{th}\,x\,\mathrm{th}\,y)\end{aligned}$$

after dividing the numerator and denominator by $\mathrm{ch}\,x\,\mathrm{ch}\,y$.

Replacing y by x we have the duplication formula for $\operatorname{th} x$:

$$\operatorname{th} 2x = 2\operatorname{th} x/(1+\operatorname{th}^2 x).$$

EXAMPLES.

$$\begin{aligned} D_x \operatorname{th} x &= D_x \operatorname{sh} x/\operatorname{ch} x - \operatorname{sh} x D_x \operatorname{ch} x/\operatorname{ch}^2 x \\ &= 1-\operatorname{sh}^2 x/\operatorname{ch}^2 x \\ &= 1-\operatorname{th}^2 x = \operatorname{sech}^2 x. \end{aligned}$$

$$D_x \coth x = D_x(1/\operatorname{th} x) = -\operatorname{sech}^2 x/\operatorname{th}^2 x = -\operatorname{cosech}^2 x.$$

$$D_x \operatorname{sech} x = -\operatorname{sh} x/\operatorname{ch}^2 x = -\operatorname{th} x \operatorname{sech} x.$$

$$D_x \operatorname{cosech} x = -\operatorname{ch} x/\operatorname{sh}^2 x = -\coth x \operatorname{cosech} x.$$

4.8. Since $e^x = \sum \frac{x^n}{n!}$, therefore $e^{-x} = \sum \frac{(-x)^n}{n!}$ and so

$$\operatorname{ch} x = \frac{1}{2}\left(\sum \frac{x^n}{n!} + \sum \frac{(-x)^n}{n!}\right) = \frac{1}{2}\sum \frac{x^n+(-x)^n}{n!} \quad \text{by 1.45,}$$

i.e.
$$\operatorname{ch} x = \sum \frac{x^{2n}}{(2n)!} = 1+\frac{x^2}{2!}+\frac{x^4}{4!}+\ldots+\frac{x^{2n}}{n!}+\ldots$$

and

$$\operatorname{sh} x = \frac{1}{2}\left(\sum \frac{x^n}{n!} - \sum \frac{(-x)^n}{n!}\right) = \frac{1}{2}\sum \frac{x^n-(-x)^n}{n!} = \sum \frac{x^{2n+1}}{(2n+1)!},$$

i.e.
$$\operatorname{sh} x = x+\frac{x^3}{3!}+\frac{x^5}{5!}+\ldots+\frac{x^{2n+1}}{(2n+1)!}+\ldots.$$

The series for $\operatorname{sh} x$ and $\operatorname{ch} x$ are absolutely convergent, being sub-series of the absolutely convergent series

$$\sum \frac{x^n}{n!} = 1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots.$$

4.9. Inverse hyperbolic functions

4.91. The derivative of $\operatorname{sh} x$ is $\operatorname{ch} x$ and $\operatorname{ch} x \geqslant 1$ for every value of x, so that by Theorem 3.83, $\operatorname{sh} x$ has a unique inverse function for all values of x, which we shall denote by $\operatorname{sh}^{-1} y$,† or $\operatorname{arg\,sh} y$ or $\operatorname{hs} y$.

Since $\operatorname{ch} x = +\sqrt{(1+\operatorname{sh}^2 x)}$ and $\operatorname{sh} x + \operatorname{ch} x = e^x$, therefore

$$\log\{\operatorname{sh} x + \sqrt{(1+\operatorname{sh}^2 x)}\} = x,$$

† The notation $\operatorname{sh}^{-1} y$ for the inverse of $\operatorname{sh} y$ is perhaps rather misleading, for though we have written $\operatorname{sh}^2 x$ for $(\operatorname{sh} x)^2$, $\operatorname{sh}^{-1} x$ does *not* stand for $(\operatorname{sh} x)^{-1} = 1/\operatorname{sh} x$. To avoid confusion we shall write $\operatorname{sh}^n x$ for $(\operatorname{sh} x)^n$ only when n is a positive integer. The origin of the notation $f^{-1}(x)$ for the inverse of $f(x)$ lies in the analogy of the equation $f^{-1}(f(x)) = x$ with the algebraic identity $f^{-1}fx = x$.

which proves that $\log\{y+\sqrt{(1+y^2)}\}$ is an inverse function of $\operatorname{sh} x$ and therefore

$$\operatorname{sh}^{-1} y = \log\{y+\sqrt{(1+y^2)}\}$$

for all values of y.

4.92. If $x \geqslant \alpha > 0$ then $\operatorname{sh} x \geqslant \operatorname{sh} \alpha > 0$, for $\operatorname{sh} x$ is steadily increasing, as may be proved directly or by observing that its derivative is $\operatorname{ch} x$ which is never less than unity. But $\operatorname{sh} x$ is the derivative of $\operatorname{ch} x$, and therefore, by 3.84, $\operatorname{ch} x$ has a unique inverse, for $x \geqslant 0$, which we shall denote by $\operatorname{ch}^{-1} y$, or $\arg \operatorname{ch} y$ or $\operatorname{hc} y$. When $x \geqslant 0$, $\operatorname{sh} x = +\sqrt{(\operatorname{ch}^2 x - 1)}$ and so $\operatorname{ch} x + \sqrt{(\operatorname{ch}^2 x - 1)} = e^x$, whence

$$\log\{\operatorname{ch} x + \sqrt{(\operatorname{ch}^2 x - 1)}\} = x, \tag{i}$$

which proves that $\log\{y+\sqrt{(y^2-1)}\}$ is the inverse of $\operatorname{ch} x$ and so

$$\operatorname{ch}^{-1} y = \log\{y+\sqrt{(y^2-1)}\}, \quad \text{where } y \geqslant 1 \text{ since } y = \operatorname{ch} x.$$

4.921. $\operatorname{sh} x$ is increasing even when x is negative, since $\operatorname{ch} x \geqslant 1$ for *all* x, and so if $\alpha > 0$ and $x \leqslant -\alpha$ then

$$\operatorname{sh} x \leqslant \operatorname{sh}(-\alpha) = -\operatorname{sh} \alpha,$$

so that $\operatorname{ch} x$ has a unique inverse when $x \leqslant 0$ which we shall denote by $\arg \operatorname{ch}^*(y)$ or $\operatorname{hc}^*(y)$. But when x is negative

$$\operatorname{sh} x = -\sqrt{(\operatorname{ch}^2 x - 1)}$$

and so $$\operatorname{ch} x - \sqrt{(\operatorname{ch}^2 x - 1)} = \operatorname{ch} x + \operatorname{sh} x = e^x,$$

whence $$\log\{\operatorname{ch} x - \sqrt{(\operatorname{ch}^2 x - 1)}\} = x,$$

proving that

$$\log\{y-\sqrt{(y^2-1)}\} = \operatorname{hc}^*(y), \quad \text{where } y \geqslant 1.$$

At the point $x = 0$, $y = 1$ we may take the inverse of $\operatorname{ch} x$ to be either $\log\{y+\sqrt{(y^2-1)}\}$ or $\log\{y-\sqrt{(y^2-1)}\}$.

Combining 4.92 and 4.921 we can say that $\operatorname{ch} x$ has a unique inverse for all values of x and that when x is positive (or zero) this inverse is $\log\{y+\sqrt{(y^2-1)}\}$ and when x is negative it is

$$\log\{y-\sqrt{(y^2-1)}\}, \quad \text{where } y = \operatorname{ch} x \geqslant 1.$$

Since $\{y-\sqrt{(y^2-1)}\}\{y+\sqrt{(y^2-1)}\} = 1$, therefore

$$\log\{y-\sqrt{(y^2-1)}\}+\log\{y+\sqrt{(y^2-1)}\} = \log 1 = 0$$

and so $\operatorname{hc}^*(y) = -\operatorname{hc}(y)$.

4.93. The derivatives of the inverse hyperbolic functions

By Theorem 3.83 $D_y \operatorname{sh}^{-1}y = 1/D_x \operatorname{sh} x = 1/\operatorname{ch} x$, where $\operatorname{sh} x = y$, and so $\operatorname{ch} x = +\sqrt{(1+y^2)}$, since $\operatorname{ch} x > 0$, and therefore

$$D_y \operatorname{sh}^{-1}y = \frac{1}{+\sqrt{(1+y^2)}}.$$

Similarly $D_y \operatorname{ch}^{-1}y = 1/D_x \operatorname{ch} x = 1/\operatorname{sh} x$, where $\operatorname{ch} x = y$ and $x > 0$, so that $\operatorname{sh} x = +\sqrt{(y^2-1)}$, and therefore

$$D_y \operatorname{ch}^{-1}y = \frac{1}{+\sqrt{(y^2-1)}}.$$

Since $\operatorname{hc}^* y = -\operatorname{ch}^{-1}y$, therefore $D_y \operatorname{hc}^* y = -\dfrac{1}{\sqrt{(y^2-1)}}$.

We can of course obtain these derivatives directly from the relations

$$\operatorname{sh}^{-1}y = \log\{y+\sqrt{(y^2+1)}\} \quad \text{and} \quad \operatorname{ch}^{-1}y = \log\{y+\sqrt{(y^2-1)}\},$$

which give

$$D_y \operatorname{sh}^{-1}y = D_y \log\{y+\sqrt{(y^2+1)}\} = \frac{1}{y+\sqrt{(y^2+1)}} D_y\{y+\sqrt{(y^2+1)}\}$$

$$= \frac{1}{y+\sqrt{(y^2+1)}}\left\{1+\frac{y}{\sqrt{(1+y^2)}}\right\} = \frac{1}{\sqrt{(1+y^2)}}$$

and

$$D_y \operatorname{ch}^{-1}y = D_y \log\{y+\sqrt{(y^2-1)}\}$$

$$= \frac{1}{y+\sqrt{(y^2-1)}}\left\{1+\frac{y}{\sqrt{(y^2-1)}}\right\} = \frac{1}{\sqrt{(y^2-1)}}.$$

V

THE CIRCULAR FUNCTIONS

ADDITION FORMULAE. DEFINITION OF π. PERIODICITY. INVERSE CIRCULAR FUNCTIONS

5. The *circular functions* $\sin x$ and $\cos x$ are defined by the equations

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots,$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots.$$

Each of the series on the right-hand side is absolutely convergent for all values of x, since $\sum \frac{|x^{2n}|}{(2n)!} = \operatorname{ch} x$ and $\sum \frac{|x^{2n+1}|}{(2n+1)!} = \operatorname{sh}|x|$; hence by 3.7

$$D_x \sin x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots = \cos x$$

and $$D_x \cos x = -x + \frac{x^3}{3!} - \frac{x^5}{5!} + \dots = -\sin x \quad \text{(see 1.6011).}$$

Thus $$D_x \sin x = \cos x \quad \text{and} \quad D_x \cos x = -\sin x.$$

5.01. $$\sin^2 x + \cos^2 x = 1.$$

For

$$D_x(\sin^2 x + \cos^2 x) = 2 \sin x D_x \sin x + 2 \cos x D_x \cos x$$
$$= 2 \sin x \cos x - 2 \cos x \sin x = 0,$$

and therefore $$\sin^2 x + \cos^2 x = \sin^2 0 + \cos^2 0 = 1$$

since $\sin 0 = 0$ and $\cos 0 = 1$.

5.02. Both $\sin x$ and $\cos x$ lie between -1 and $+1$; for $\sin^2 x$ and $\cos^2 x$ are necessarily positive and so

$$0 \leqslant \sin^2 x = 1 - \cos^2 x \leqslant 1$$

and $$0 \leqslant \cos^2 x = 1 - \sin^2 x \leqslant 1$$

and therefore both $\sin^2 x$ and $\cos^2 x$ lie between 0 and 1, which proves that $\sin x$ and $\cos x$ lie between -1 and $+1$.

5.1. Addition formulae

$$\sin(x+y) = \sin x \cos y + \cos x \sin y,$$
$$\cos(x+y) = \cos x \cos y - \sin x \sin y.$$

Write $S(x)$ and $C(x)$ for the right-hand sides of these equations, so that $D_x S(x) = C(x)$ and $D_x C(x) = -S(x)$. Then

$$D_x\{S(x)\cos(x+y)-C(x)\sin(x+y)\} = C(x)\cos(x+y)-$$
$$-S(x)\sin(x+y)-C(x)\cos(x+y)+S(x)\sin(x+y) = 0,$$

and therefore

$$S(x)\cos(x+y)-C(x)\sin(x+y) = S(0)\cos y-C(0)\sin y = 0$$

since $\quad S(0) = \sin 0\cos y+\cos 0\sin y = \sin y$

and $\quad C(0) = \cos 0\cos y-\sin 0\sin y = \cos y.$

Thus $\quad S(x)\cos(x+y)-C(x)\sin(x+y) = 0 \qquad \text{(i)}$

and similarly

$$S(x)\sin(x+y)+C(x)\cos(x+y) = 1. \qquad \text{(ii)}$$

Multiply equation (i) by $\cos(x+y)$ and (ii) by $\sin(x+y)$ and add, then since $\sin^2(x+y)+\cos^2(x+y) = 1$, we have

$$S(x) = \sin(x+y).$$

Multiply (ii) by $\cos(x+y)$ and (i) by $\sin(x+y)$ and subtract and we find

$$C(x) = \cos(x+y),$$

which completes the proof.

5.11. Since $\sin x$ is a sum of odd powers of x, $\sin(-x) = -\sin x$, and since $\cos x$ is a sum of even powers, $\cos(-x) = \cos x$; hence writing $-y$ for y in 5.1 we have

$$\sin(x-y) = \sin x\cos y-\cos x\sin y,$$
$$\cos(x-y) = \cos x\cos y+\sin x\sin y.$$

5.12. Duplication formulae

Write x for y in 5.1 and we have

$$\sin 2x = 2\sin x\cos y,$$
$$\cos 2x = \cos^2 x-\sin^2 x$$
$$= 2\cos^2 x-1, \quad \text{by 5.01},$$
$$= 1-2\sin^2 x.$$

5.13. If

$$\cos_n(x) = 1-x^2/2!+x^4/4!-...+(-1)^n x^{2n}/(2n)!$$

and $\quad \sin_n(x) = x-x^3/3!+x^5/5!-...+(-1)^n x^{2n+1}/(2n+1)!$

then for any x and n, $\cos x$ lies between $\cos_n(x)$ and $\cos_{n+1}(x)$ and $\sin x$ lies between $\sin_n(x)$ and $\sin_{n+1}(x)$.

For $D_x \sin_n(x) = \cos_n(x)$ and $D_x \cos_n(x) = -\sin_{n-1}(x)$, and so if $P_n(x) = \cos_n(x) - \cos x$ and $Q_n(x) = \sin_n(x) - \sin x$ then

$$P'_n(x) = -Q_{n-1}(x), \qquad Q'_n(x) = P_n(x), \qquad P_n(0) = Q_n(0) = 0.$$

Hence, for $x > 0$, *if* $P_{2n}(x) > 0$ then $Q_{2n}(x)$ is increasing and so $Q_{2n}(x) > 0$, and therefore $P_{2n+1}(x) < 0$; and if $P_{2n}(x) < 0$ then $Q_{2n}(x) < 0$ and so $P_{2n+1}(x) > 0$. Similarly, if $P_{2n+1}(x) > 0$ then $P_{2n+2}(x) < 0$, and if $P_{2n+1}(x) < 0$ then $P_{2n+2}(x) > 0$. But $P_0(x) > 0$, and therefore, in turn, $P_1(x) < 0$, $P_2(x) > 0$, $P_3(x) < 0$, and so on, so that $P_n(x)$ and $P_{n+1}(x)$ have opposite signs. Similarly $Q_n(x)$ and $Q_{n+1}(x)$ have opposite signs. If $x < 0$, $P_n(x) = P_n(-x)$ and $Q_n(x) = -Q_n(-x)$, and so also when $x < 0$, $P_n(x)$ and $P_{n+1}(x)$ have opposite signs, and $Q_n(x)$ and $Q_{n+1}(x)$ have opposite signs. Hence $\cos x$ lies between $\cos_n(x)$ and $\cos_{n+1}(x)$ and $\sin x$ lies between $\sin_n(x)$ and $\sin_{n+1}(x)$.

5.2. We prove next that the equation $\cos x = 0$ has a solution lying between 1·5 and 1·6.

By 5.13, $\cos\frac{3}{2}$ lies between $1-\frac{(\frac{3}{2})^2}{2!}+\frac{(\frac{3}{2})^4}{4!}$ and $1-\frac{(\frac{3}{2})^2}{2!}+\frac{(\frac{3}{2})^4}{4!}-\frac{(\frac{3}{2})^6}{6!}$ and so $\cos\frac{3}{2}$ *exceeds*

$$1-\frac{(\frac{3}{2})^2}{2!}+\frac{(\frac{3}{2})^4}{4!}-\frac{(\frac{3}{2})^6}{6!} = 1-\frac{9}{8}+\frac{3^4}{2^4.4!}\left(1-\frac{3}{40}\right) > 1-\frac{9}{8}+\frac{1}{8} = 0,$$

since $$1-\tfrac{3}{40} > \tfrac{16}{27};$$

thus $\cos\frac{3}{2} > 0$.

Similarly $\cos\frac{8}{5}$ is *less* than

$$1-\frac{(\frac{8}{5})^2}{2!}+\frac{(\frac{8}{5})^4}{4!} = 1-\frac{32}{25}\left(1-\frac{16}{75}\right) < 0$$

because $$32\times 59 = 1888 > 1875 = 25\times 75;$$

thus $\cos\frac{8}{5} < 0$.

But $\cos x$ is continuous, and therefore, by 2.4, there is a point between $\frac{3}{2} = 1{\cdot}5$ and $\frac{8}{5} = 1{\cdot}6$, where $\cos x = 0$.

If x lies between 1 and $\frac{8}{5}$, then $x^2+x+1 < 3!$ so that

$$(x^3-1) < 3!(x-1),$$

i.e. $x-\dfrac{x^3}{3!} > 1-\dfrac{1}{3!}$; but $\sin x > x-\dfrac{x^3}{3!}$ and therefore

$$\sin x > 1-\frac{1}{3!} = \frac{5}{6}.$$

Since the derivative of $\cos x = -\sin x < -\frac{5}{6}$ between 1 and $\frac{8}{5}$, therefore $\cos x$ steadily decreases as x passes from 1 to $\frac{8}{5}$, and therefore $\cos x$ takes each of its values *once only* between 1 and $\frac{8}{5}$. In particular $\cos x = 0$ *once only* between 1 and $\frac{8}{5}$; we denote this unique value of x by $\frac{1}{2}\pi$ (read 'half pi'), so that $\cos\frac{1}{2}\pi = 0$; we have seen that $\frac{1}{2}\pi$ lies between 1·5 and 1·6, so that π lies between 3 and 3·2. The method of 2.4 enables us to calculate the number $\frac{1}{2}\pi$ to as many decimal places as we please, but we shall later give less laborious methods by which this calculation may be effected. To 25 places of decimals

$$\pi = 3{\cdot}14159\ 26535\ 89793\ 23846\ 26433.$$

Since $\sin x > x - \frac{x^3}{6} > 0$, provided $0 < x < \sqrt{6}$, therefore $\cos x$ is steadily decreasing between 0 and $\frac{1}{2}\pi$, and so $\frac{1}{2}\pi$ is the smallest positive root of the equation $\cos x = 0$.

5.21. Periodicity of the circular functions

Since $\cos\frac{1}{2}\pi = 0$, therefore by 5.02, $\sin^2\frac{1}{2}\pi = 1$; but $\sin x >$ when x lies between 1 and $\frac{8}{5}$, and $\frac{1}{2}\pi$ falls in this range, and therefore

$$\sin\tfrac{1}{2}\pi = +1.$$

By 5.12

$$\cos\pi = \cos^2\tfrac{1}{2}\pi - \sin^2\tfrac{1}{2}\pi = -1, \quad \text{i.e. } \cos\pi = -1,$$

and

$$\sin\pi = 2\sin\tfrac{1}{2}\pi\cos\tfrac{1}{2}\pi = 0, \quad \text{i.e. } \sin\pi = 0.$$

Hence by 5.1

$$\cos(\pi+x) = -\cos x,$$
$$\sin(\pi+x) = -\sin x.$$

Furthermore

$$\cos 2\pi = \cos^2\pi - \sin^2\pi = 1, \qquad \sin 2\pi = 2\sin\pi\cos\pi =$$

and so

$$\cos(x+2\pi) = \cos x, \qquad \sin(x+2\pi) = \sin x. \qquad ($$

Thus $\sin x$ *and* $\cos x$ *are periodic functions, of period* 2π

Replacing x by $x+2\pi$ in (i) we find that

$$\cos(x+4\pi) = \cos x, \qquad \sin(x+4\pi) = \sin x;$$

repeating the transformation, we see that

$$\cos(x+6\pi) = \cos x, \qquad \sin(x+6\pi) = \sin x,$$

and so on up to any positive integer n,

$$\cos(x+2n\pi) = \cos x, \qquad \sin(x+2n\pi) = \sin x$$

Replacing x by $-x$ and remembering that

$$\cos(-x) = \cos x, \qquad \sin(-x) = -\sin x \quad ,$$

we have

$$\cos(x-2n\pi) = \cos x, \qquad \sin(x-2n\pi) = \sin x,$$

and so $$\cos(x+2m\pi) = \cos x, \qquad \sin(x+2m\pi) = \sin x$$

whether m be positive or negative.

All that remains to complete the identification of the functions defined in 5 with the familiar trigonometric functions is to prove the fundamental theorem of trigonometry (or one may say, the *trigonometric* definition of $\sin x$) that in a unit circle an arc of length $2x$ stands on a chord of length $2\sin x$. As this is but a special case of general theorems concerning the length of a curved line we delay the proof till the appropriate point is reached in the general development of the subject.

5.3. The subsidiary circular functions $\tan x$, $\cot x$, $\sec x$, $\operatorname{cosec} x$ have the familiar definitions, viz.

$$\tan x = \sin x/\cos x, \qquad \cot x = 1/\tan x,$$

$$\sec x = 1/\cos x, \qquad \operatorname{cosec} x = 1/\sin x.$$

Hence

$$D_x \tan x = (\cos^2 x+\sin^2 x)/\cos^2 x = \sec^2 x,$$

$$D_x \cot x = -\sec^2 x/\tan^2 x = -1/\sin^2 x = -\operatorname{cosec}^2 x,$$

$$D_x \sec x = -(-\sin x)/\cos^2 x = \sec x \tan x,$$

$$D_x \operatorname{cosec} x = -\cos x/\sin^2 x = -\operatorname{cosec} x \cot x.$$

5.31. From 5.01, dividing in turn by $\cos^2 x$ and $\sin^2 x$, we find

$$\sec^2 x = 1+\tan^2 x \quad \text{and} \quad \operatorname{cosec}^2 x = 1+\cot^2 x.$$

Furthermore, from 5.1,

$$\begin{aligned}\tan(x+y) &= \sin(x+y)/\cos(x+y)\\ &= (\sin x \cos y + \cos x \sin y)/(\cos x \cos y - \sin x \sin y)\\ &= (\tan x + \tan y)/(1-\tan x \tan y)\end{aligned}$$

after dividing numerator and denominator by $\cos x \cos y$.

Since $\sin(-x) = -\sin x$, $\cos(-x) = \cos x$, therefore

$$\tan(-x) = -\tan x;$$

hence replacing y by $-y$ in the formula for $\tan(x+y)$, we find

$$\tan(x-y) = (\tan x - \tan y)/(1+\tan x \tan y).$$

5.4. α is a positive number, however small; then $\sin x$ is *positive* when x lies in the interval $(\alpha, \pi-\alpha)$ and *negative* when x lies in $(\pi+\alpha, 2\pi-\alpha)$, i.e. $\sin x$ is positive in the *open* interval $[0, \pi]$ and negative in the *open* interval $[\pi, 2\pi]$.

We have already observed that if $0 < x \leqslant 2$ then $\sin x > x - \dfrac{x^3}{3!}$. But if $0 < \alpha < \frac{1}{2}$ and $\alpha \leqslant x \leqslant 2$, then $x^2+\alpha x+\alpha^2 < 6$ and $x-\alpha \geqslant 0$, and so $x^3-\alpha^3 \leqslant 6(x-\alpha)$, whence $x-\dfrac{x^3}{3!} \geqslant \alpha-\dfrac{\alpha^3}{3!} \geqslant \frac{5}{6}\alpha$, as $\alpha^3 < \alpha$. Thus if $\alpha \leqslant x \leqslant 2$, then $\sin x \geqslant \frac{5}{6}\alpha$. But $1 < \frac{1}{2}\pi < 2$ and so $\sin x \geqslant \frac{5}{6}\alpha$ for any x in $(\alpha, \frac{1}{2}\pi)$.

Since $\sin(\pi-x) = \sin x$ and if x lies in $(\frac{1}{2}\pi, \pi-\alpha)$ then $\pi-x$ lies in $(\alpha, \frac{1}{2}\pi)$ therefore $\sin x \geqslant \frac{5}{6}\alpha$ also when x lies in $(\frac{1}{2}\pi, \pi-\alpha)$. Thus if $0 < \alpha \leqslant x \leqslant \pi-\alpha$, $\sin x \geqslant \frac{5}{6}\alpha$ and this is true however small α may be.

Moreover, $\sin(x+\pi) = -\sin x$, and if x lies in $(\alpha, \pi-\alpha)$ then $x+\pi$ lies in $(\pi+\alpha, 2\pi-\alpha)$ and therefore at any point of the interval $(\pi+\alpha, 2\pi-\alpha)$, $\sin x \leqslant -\frac{5}{6}\alpha$.

5.41. $\cos x$ is *positive* in the open interval $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ and *negative* in the open interval $[\frac{1}{2}\pi, \frac{3}{2}\pi]$.

This follows directly from § 5.4, for $\cos x = \sin(x+\frac{1}{2}\pi)$, and when x lies in $(-\frac{1}{2}\pi+\alpha, \frac{1}{2}\pi-\alpha)$, then $x+\frac{1}{2}\pi$ lies in $(\alpha, \pi-\alpha)$ and therefore $\cos x \geqslant \frac{5}{6}\alpha$; and when x lies in $(\frac{1}{2}\pi+\alpha, \frac{3}{2}\pi-\alpha)$ then $x+\frac{1}{2}\pi$ lies in $(\pi+\alpha, 2\pi-\alpha)$ and so $\cos x \leqslant -\frac{5}{6}\alpha$.

5.42. The periodicity of $\sin x$ and $\cos x$ enables us to deduce from 5.4 and 5.41 that $\sin x$ and $\cos x$ are positive in $[2r\pi, (2r+1)\pi]$ and $[(2r-\frac{1}{2})\pi, (2r+\frac{1}{2})\pi]$ respectively and negative in

$$[(2r+1)\pi, 2(r+1)\pi] \quad \text{and} \quad [(2r+\tfrac{1}{2})\pi, (2r+\tfrac{3}{2})\pi]$$

respectively.

5.5. Inverse circular functions

The derivative of $\cos x$ is $-\sin x$ for all x and $-\sin x \leqslant -\frac{5}{6}\alpha$, if x lies in the interval $(\alpha, \pi-\alpha)$. Hence by 3.84, $\cos x$ has a unique *continuous* inverse in $(0, \pi)$ which we shall denote by $\cos^{-1}y$ or $\operatorname{arc}\cos y$, where since $y = \cos x$, the values of y lie between ± 1. It is very important to keep clearly in mind the *interval* in which $\cos^{-1}y$ is the inverse of $\cos x$, i.e. the range of values of $\cos^{-1}y$, viz. from 0 to π.

5.501. In the interval $(2r\pi, (2r+1)\pi)$ both $\sin x$ and $\cos x$ have the same values as those taken in the interval $(0, \pi)$ and so $\cos x$ has a unique inverse in the interval $(2r\pi, (2r+1)\pi)$, which we shall denote by $\operatorname{arc}\cos_r y$, the suffix r corresponding to the r in the specification of the interval.

If x lies in the interval $(0, \pi)$ then $x+2r\pi$ lies in $(2r\pi, (2r+1)\pi)$ and therefore

$$\operatorname{arc}\cos(\cos x) = x$$

and
$$\operatorname{arc}\cos_r(\cos(x+2r\pi)) = x+2r\pi.$$

But $\cos(x+2r\pi) = \cos x$, and so if $y = \cos x$ then

$$\operatorname{arc}\cos_r y = \operatorname{arc}\cos y+2r\pi,$$

which is the fundamental relation between $\operatorname{arc}\cos_r y$ and $\operatorname{arc}\cos y$.

5.502. If x lies in the interval $(\pi+\alpha, 2\pi-\alpha)$ then $x-\pi$ lies in $(\alpha, \pi-\alpha)$ and so, since $\sin x = -\sin(x-\pi)$, it follows that

$$\sin x < -\tfrac{5}{8}\alpha.$$

Thus $\cos x$ has a unique inverse in the interval $(\pi, 2\pi)$ which we shall denote by $\operatorname{arc}\cos^* y$. Similarly $\cos x$ has a unique inverse in $((2r+1)\pi, 2(r+1)\pi)$ which we denote by $\operatorname{arc}\cos_r^* y$.

If x lies between 0 and π then $2\pi-x$ lies between π and 2π and so

$$\operatorname{arc}\cos(\cos x) = x$$

and
$$\operatorname{arc}\cos^*(\cos(2\pi-x)) = 2\pi-x.$$

But $\cos(2\pi-x) = \cos x = y$ and therefore

$$\operatorname{arc}\cos^* y = 2\pi-\operatorname{arc}\cos y.$$

If x lies in $(0, \pi)$ then $2(r+1)\pi-x$ lies in $((2r+1)\pi, 2(r+1)\pi)$ and so, since $\cos(2(r+1)\pi-x) = \cos x$, we have

$$\operatorname{arc}\cos_r^* y = 2(r+1)\pi-\operatorname{arc}\cos y.$$

5.503. By Theorem 3.832 the derivative of $\operatorname{arc}\cos y$ is

$$1/D_x\cos x = -1/\sin x,$$

where $\cos x = y$, and $0 < x < \pi$. But $\sin^2 x = 1-\cos^2 x$ and so, since $\sin x$ is positive when $0 < x < \pi$,

$$\sin x = +\sqrt{(1-\cos^2 x)} = +\sqrt{(1-y^2)}.$$

Therefore
$$D_y \operatorname{arc}\cos y = -\frac{1}{\sqrt{(1-y^2)}}.$$

Observe that when $x = 0$ or π, $y = \pm 1$ and $1-y^2 = 0$ so that $\operatorname{arc}\cos y$ has no derivative at these points.

From the relation $\operatorname{arc\,cos}_r y = \operatorname{arc\,cos} y + 2r\pi$ it follows that

$$D_y \operatorname{arc\,cos}_r y = -\frac{1}{\sqrt{(1-y^2)}}$$

and from

$$\operatorname{arc\,cos}_r^* y = 2(r+1)\pi - \cos^{-1} y$$

that

$$D_y \operatorname{arc\,cos}_r^* y = \frac{1}{\sqrt{(1-y^2)}}.$$

In particular the inverse of $\cos x$ in $(\pi,\ 2\pi)$ has the derivative $1/\sqrt{(1-y^2)}$.

5.504. The determination of the inverse functions of $\sin x$ may be carried out in the same way. Since $\cos x > \frac{5}{8}\alpha$ when

$$-\tfrac{1}{2}\pi + \alpha \leqslant x \leqslant \tfrac{1}{2}\pi - \alpha,$$

we find that in the interval $(-\frac{1}{2}\pi,\ \frac{1}{2}\pi)$ the function $\sin x$ has a unique continuous inverse $\operatorname{arc\,sin} y$, of which the derivative is

$$1/D_x \sin x = +1/\sqrt{(1-y^2)}.$$

In the interval $((2r-\frac{1}{2})\pi, (2r+\frac{1}{2})\pi)$ $\sin x$ has the unique inverse $\operatorname{arc\,sin}_r y$ such that $\operatorname{arc\,sin}_r y = \operatorname{arc\,sin} y + 2r\pi$ and

$$D_y \operatorname{arc\,sin}_r y = +1/\sqrt{(1-y^2)}.$$

In the interval $((2r+\frac{1}{2})\pi, (2r+\frac{3}{2})\pi)$ $\sin x$ has the unique inverse $\operatorname{arc\,sin}_r^* y$ such that $\operatorname{arc\,sin}_r^* y = (2r+1)\pi - \operatorname{arc\,sin} y$ and

$$D_y \operatorname{arc\,sin}_r^* y = -1/\sqrt{(1-y^2)}.$$

5.51. The inverse of $\tan x$

Since $\cos x > \frac{5}{8}\alpha$ in the interval $(-\frac{1}{2}\pi+\alpha,\ \frac{1}{2}\pi-\alpha)$, therefore $\tan x$ is differentiable in this interval, for any $\alpha > 0$. Furthermore, if N is any positive number, however great, we may take

$$\alpha = 1/2N \quad \text{and so} \quad \tan(\tfrac{1}{2}\pi-\alpha) = \frac{\cos\alpha}{\sin\alpha} > \frac{1-\frac{1}{2}\alpha^2}{\alpha} = \frac{1}{\alpha} - \tfrac{1}{2}\alpha > N,$$

and furthermore $\tan(-\frac{1}{2}\pi+\alpha) = -\tan(\frac{1}{2}\pi-\alpha) < -N$; since $\tan x$ is continuous in $(-\frac{1}{2}\pi+\alpha,\ \frac{1}{2}\pi-\alpha)$ and exceeds $+N$ and is exceeded by $-N$ in this interval (if $\alpha = 1/2N$) therefore $\tan x$ takes both the values $\pm N$ in this interval, be N as great as we please. Hence, by 3.841, $\tan x$ has a unique inverse in the open interval $[-\frac{1}{2}\pi,\ \frac{1}{2}\pi]$, which we shall denote by $\tan^{-1} y$ or $\operatorname{arc\,tan} y$; $\tan^{-1} y$ is defined for *all* values of y and its derivative is

$$1/\{\sec(\tan^{-1} y)\}^2 = 1/[1+\{\tan(\tan^{-1} y)\}^2] = 1/(1+y^2),$$

also for all values of y.

5.52.* When x lies in the interval $((2r-1)\frac{1}{2}\pi+\alpha,\ (2r+1)\frac{1}{2}\pi-\alpha)$ then $x-r\pi$ lies in $(-\frac{1}{2}\pi+\alpha,\ \frac{1}{2}\pi-\alpha)$ and so, since

$$\cos x = \pm\cos(x-r\pi),$$

according as r is even or odd, we have $\cos x > \frac{5}{8}\alpha$ or $\cos x < -\frac{5}{8}\alpha$; in either case $\tan x$ is differentiable in the interval

$$((2r-1)\tfrac{1}{2}\pi+\alpha,\ (2r+1)\tfrac{1}{2}\pi-\alpha)$$

and its derivative is $\sec^2 x \geqslant 1$. Hence, as in 5.51, $\tan x$ has a unique inverse in the open interval $[(2r-1)\frac{1}{2}\pi, (2r+1)\frac{1}{2}\pi]$, which we shall denote by $\operatorname{arc\,tan}_r y$. The inverse function $\operatorname{arc\,tan}_r y$ is defined for all values of y.

If x lies in $[-\frac{1}{2}\pi,\ \frac{1}{2}\pi]$ then $x+r\pi$ lies in $[(2r-1)\frac{1}{2}\pi, (2r+1)\frac{1}{2}\pi]$ and therefore

$$\operatorname{arc\,tan}(\tan x) = x$$

and

$$\operatorname{arc\,tan}_r(\tan(x+r\pi)) = x+r\pi,$$

whence, since $\tan(x+r\pi) = \tan x = y$, we have

$$\operatorname{arc\,tan}_r y = \operatorname{arc\,tan} y+r\pi$$

and therefore

$$D_y \operatorname{arc\,tan}_r y = D_y \operatorname{arc\,tan} y = 1/(1+y^2)$$

for all values of y.

5.53. $\operatorname{arc\,tan} x+\operatorname{arc\,tan} y = \operatorname{arc\,tan}\dfrac{x+y}{1-xy}$ provided $xy <$

For $$D_x(\operatorname{arc\,tan} x + \operatorname{arc\,tan} y) = 1/(1+x^2)$$

and

$$D_x(\operatorname{arc\,tan}(x+y)/(1-xy)) = \frac{(1+y^2)}{(1-xy)^2}\bigg/\frac{(1+y^2)(1+x^2)}{(1-xy)^2}$$
$$= 1/(1+x^2) \quad \textit{provided } xy \neq 1.$$

When $x = 0$,

$$\operatorname{arc\,tan} x + \operatorname{arc\,tan} y = \operatorname{arc\,tan} y = \operatorname{arc\,tan}(x+y)/(1-xy),$$

for all y, and therefore $\operatorname{arc\,tan} x + \operatorname{arc\,tan} y = \operatorname{arc\,tan}\dfrac{x+y}{1-xy}$ in any interval of values of x which contains $x = 0$, provided $xy \neq 1$; hence the equality holds for all x and y such that $xy < 1$, for the range of values of x for which $xy > 1$ does not contain the point $x = 0$.

VI
PARTIAL FRACTIONS

6. In this chapter we shall show how a rational function $P(x)/Q(x)$, where $P(x)$, $Q(x)$ are polynomials, may be expressed as a sum of terms each of which is either a polynomial or of the form $A/(x+a)^p$ or of the form $(Ax+B)/\{(x+\lambda)^2+a\}^p$. The expression of $P(x)/Q(x)$ in this way is called the resolution of $P(x)/Q(x)$ into *partial fractions.*

We shall not prove a general theorem on the resolution of a rational function into partial fractions but shall content ourselves with a description of a number of devices by which this resolution may be effected with a minimum of calculation in the cases most commonly met with in practice. We shall always suppose that $P(x)/Q(x)$ is in its lowest terms and that the degree of $P(x)$ is less than that of $Q(x)$, for if this is not the case we may divide $Q(x)$ into $P(x)$, obtaining a polynomial quotient $P^*(x)$ and a remainder $R(x)$, of degree less than $Q(x)$, such that

$$P(x)/Q(x) = P^*(x)+R(x)/Q(x).$$

CASE 1. $Q(x)$ contains a simple linear factor $x-a$ and

$$Q(x) = (x-a)Q^*(x).$$

We choose a constant A so that $P(x)-AQ^*(x)$ has the factor $x-a$; this requires $P(a)-AQ^*(a) = 0$, i.e. $A = P(a)/Q^*(a)$, since $Q^*(a) \neq 0$ else $x-a$ would be a repeated factor of $Q(x)$. Thus $P(x)-P(a)Q^*(x)/Q^*(a)$ is divisible by $(x-a)$; let the quotient be $P^*(x)$ so that

$$P(x)-P(a)Q^*(x)/Q^*(a) = (x-a)P^*(x),$$

$$\therefore \quad \frac{P(x)}{Q(x)} = \frac{P(a)}{Q^*(a)}\cdot\frac{1}{x-a}+\frac{P^*(x)}{Q^*(x)}.$$

Thus $\dfrac{P(a)}{Q^*(a)}\cdot\dfrac{1}{x-a}$ is the partial fraction corresponding to the factor $x-a$.

If $Q(x)$ is a product of n linear factors $x-a_r$, $r = 1, 2,..., n$, and if $Q_r(x) = Q(x)/(x-a_r)$ then

$$\frac{P(x)}{Q(x)} = \frac{P(a_1)}{Q_1(a_1)}\cdot\frac{1}{x-a_1}+\frac{P(a_2)}{Q_2(a_2)}\cdot\frac{1}{x-a_2}+\dots+\frac{P(a_n)}{Q_n(a_n)}\cdot\frac{1}{x-a_n}.$$

Proof. The polynomial

$$R(x) = P(x)-A_1 Q_1(x)-A_2 Q_2(x)-...-A_n Q_n(x),$$

of degree less than n, has the factor $x-a_r$ if $R(a_r) = 0$, that is if $P(a_r) = A_r Q_r(a_r)$, since

$$Q_s(a_r) = Q(a_r)/(a_r-a_s) = 0, \quad \text{if } s \neq r.$$

Hence if $A_r = P(a_r)/Q_r(a_r)$ for all r, between 1 and n, then

$$R(x) = \lambda(x-a_1)(x-a_2)...(x-a_n),$$

where λ is constant; equating the coefficients of x^n on both sides, $\lambda = 0$, and so $R(x)$ is identically zero.

Dividing both sides of the identity

$$P(x) = A_1 Q_1(x)+...+A_n Q_n(x)$$

by $Q(x)$ the result stated follows.

Example.

$$\frac{x^2}{(x+2)(x-1)(x-3)} \quad \frac{(-2)^2}{-3.-5}\cdot\frac{1}{x+2}+\frac{1^2}{3.-2}\cdot\frac{1}{x-1}+\frac{3^2}{5.2}\cdot\frac{1}{x-3}$$

$$-\frac{4}{15}\cdot\frac{1}{x+2}\cdot\ \frac{1}{6}\cdot\frac{1}{x-1}+\frac{9}{10}\cdot\frac{1}{x-3}.$$

To find the partial fraction corresponding to a simple factor $x-a$ it is not necessary to find $Q^*(x)$ the quotient of $Q(x)$ divided by $x-a$, for if

$$Q(x) = (x-a)Q^*(x) \quad \text{then} \quad Q'(x) = Q^*(x)+(x-a)Q^{*\prime}(x)$$

and therefore $Q'(a) = Q^*(a)$, whence it follows that the partial fraction corresponding to the factor $(x-a)$ is

$$\frac{P(a)}{Q'(a)}\cdot\frac{1}{x-a}.$$

Example. x^3+x-2 has the factor $x-1$, and its derivative is $3x^2+1$. Therefore the partial fraction of $\dfrac{x+3}{x^3+x-2}$ corresponding to this factor is

$$\frac{1+3}{3.1^2+1}\cdot\frac{1}{x-1} = \frac{1}{x-1}.$$

Case 2. $Q(x)$ contains the factor $(x-a)^p$ and

$$Q(x) = (x-a)^p Q^*(x).$$

Write $(a+h)$ for x and express $Q^*(a+h)$ and $P(a+h)$ in ascending powers of h. Divide $P(a+h)$ by $Q^*(a+h)$ until the remainder is divisible by h^p; let the quotient be

$$A_0+A_1h+A_2h^2+...+A_{p-1}h^{p-1}$$

and the remainder $h^pR(h)$ so that

$$\frac{P(a+h)}{Q^*(a+h)} = A_0+A_1h+A_2h^2+...+A_{p-1}h^{p-1}+\frac{h^pR(h)}{Q^*(a+h)}$$

and therefore

$$\frac{P(a+h)}{h^pQ^*(a+h)} = \frac{A_0}{h^p}+\frac{A_1}{h^{p-1}}+...+\frac{A_{p-1}}{h}+\frac{R(h)}{Q^*(a+h)},$$

i.e.

$$\frac{P(x)}{(x-a)^pQ^*(x)} = \frac{A_0}{(x-a)^p}+\frac{A_1}{(x-a)^{p-1}}+...+\frac{A_{p-1}}{x-a}+\frac{R(x-a)}{Q^*(x)}.$$

Example. To express $\dfrac{x^3+x+1}{(x-1)^3(x^2+2)}$ in partial fractions.

Write $x-1 = h$, then

$$1+x+x^3 = 3+4h+3h^2+h^3, \qquad 2+x^2 = 3+2h+h^2.$$

$$\begin{array}{r|l}
3+2h+h^2 \,) & 3+4h+3h^2+h^3 \,(\, 1+\tfrac{2}{3}h+\tfrac{2}{9}h^2 \\
 & \quad 2h+2h^2+h^3 \\
 & \qquad \tfrac{2}{3}h^2+\tfrac{1}{3}h^3 \\
 & \qquad\quad -\tfrac{1}{9}h^3-\tfrac{2}{9}h^4
\end{array}$$

Therefore

$$\frac{x^3+x+1}{(x-1)^3(x^2+2)} = \frac{1}{h^3}+\frac{2}{3}\frac{1}{h^2}+\frac{2}{9}\frac{1}{h}-\frac{1}{9}\frac{2h+1}{3+2h+h^2}$$

$$= \frac{1}{(x-1)^3}+\frac{2}{3(x-1)^2}+\frac{2}{9(x-1)}-\frac{2x-1}{9(x^2+2)}.$$

The accuracy of the division may be tested by giving x particular values, for instance $x = 0$ and $x = 2$.

Case 3. $Q(x)$ is a product of binomial quadratic factors, i.e. quadratic factors which contain no term in x. Separate the even and odd powers of x in $P(x)$ so that $P(x)$ is expressed in the form $A(x^2)+xB(x^2)$. Consider separately $A(x^2)/Q(x)$ and $B(x^2)/Q(x)$. Write y for x^2 and the resulting expressions fall into one of the previous forms.

Example.

$$\frac{x^3+x^2+3x+4}{(x^2+1)(x^2+2)^2}$$

$$=\frac{x^2+4}{(x^2+1)(x^2+2)^2}+x\frac{x^2+3}{(x^2+1)(x^2+2)^2}$$

$$=\frac{y+4}{(y+1)(y+2)^2}+x\frac{y+3}{(y+1)(y+2)^2}, \quad \text{writing } y \text{ for } x^2,$$

$$=\frac{3}{y+1}-\frac{3}{y+2}-\frac{2}{(y+2)^2}+x\left\{\frac{2}{y+1}-\frac{2}{y+2}-\frac{1}{(y+2)^2}\right\}$$

$$=\frac{2x+3}{x^2+1}-\frac{2x+3}{x^2+2}-\frac{x+2}{(x^2+2)^2}.$$

CASE 4. $Q(x)$ is a product of linear and binomial quadratic factors.

For each linear factor $x+a$ multiply $P(x)$ and $Q(x)$ by $x-a$ and the problem is reduced to that in Case 3.

Example.

$$\frac{x^2+1}{(x+1)(x^2+2)^2}=\frac{(x-1)(x^2+1)}{(x^2-1)(x^2+2)^2}$$

$$=(x-1)\left\{\frac{y+1}{(y-1)(y+2)^2}\right\}, \quad \text{writing } y \text{ for } x^2,$$

$$=(x-1)\left\{\frac{1}{3(y+2)^2}+\frac{}{9(y-1)}-\frac{}{9(y+2)}\right\}$$

$$=\frac{2(x-1)}{9(x^2-1)}-\frac{2(x-1)}{9(x^2+2)}+\frac{x-1}{3(x^2+2)^2}$$

$$=\frac{2}{9(x+1)}-\frac{2(x-1)}{9(x^2+2)}+\frac{x-1}{3(x^2+2)^2}.$$

CASE 5. $Q(x)$ contains trinomial quadratic factors.

Example (*a*). To express $\dfrac{x^3+3x+11}{(x^2+x+1)(x^2+2x+3)}$ in partial fractions.

Since $x^3+3x+11=(x^2+x+1)(x-1)+3x+12$ the given expression is equal to

$$\frac{x-1}{x^2+2x+3}+\frac{3x+12}{(x^2+x+1)(x^2+2x+3)}.$$

Now
$$\frac{1}{x^2+x+1}+\frac{1}{x^2+2x+3}=\frac{2x^2+3x+4}{(x^2+x+1)(x^2+2x+3)} \qquad \text{(i)}$$
$$\frac{1}{x^2+x+1}-\frac{1}{x^2+2x+3}=\frac{x+2}{(x^2+x+1)(x^2+2x+3)} \qquad \text{(ii)}$$

Next we express $3x+12$ in the form $A(2x^2+3x+4)+B(x+2)$

Since $$2x(x+2)-(2x^2+3x+4) = x-4$$
$$(x+2)-(x-4) = 6,$$
$$2(x+2)+(x-4) = 3x,$$
therefore
$$3x+12 = 4(x+2)-(x-4) = (4-2x)(x+2)+(2x^2+3x+4)$$
Add equation (i) to $(4-2x)$ times equation (ii) and we find
$$\frac{5-2x}{x^2+x+1}+\frac{2x-3}{x^2+2x+3}=\frac{3x+12}{(x^2+x+1)(x^2+2x+3)} \qquad \text{(iii)}$$
Hence the expression is equal to
$$\frac{x-1}{x^2+2x+3}+\frac{5-2x}{x^2+x+1}+\frac{2x-3}{x^2+2x+3}=\frac{5-2x}{x^2+x+1}+\frac{3x-4}{x^2+2x+3}$$
which are the required partial fractions.

Example (*b*). To express $\dfrac{3x+12}{(x^2+x+1)(x^2+2x+3)(x^2+3x+4)}$ in partial fractions.

From equation (iii) above (Example (*a*))
$$\frac{3x+12}{(x^2+x+1)(x^2+2x+3)(x^2+3x+4)}$$
$$=\frac{5-2x}{(x^2+x+1)(x^2+3x+4)}+\frac{2x-3}{(x^2+2x+3)(x^2+3x+4)}. \qquad \text{(iv)}$$
Sum and difference of x^2+x+1, x^2+3x+4 is $2x^2+4x+5$ and $2x+3$.

Since $(2x^2+4x+5)-x(2x+3) = x+5$ and $2(x+5)-(2x+3) = 7$, $5(2x+3)-3(x+5) = 7x$, therefore $16(x+5)-15(2x+3) = 7(5-2x)$ and so $16(2x^2+4x+5)-(16x+15)= 7(5-2x)$. Hence
$$\frac{1-16x}{x^2+x+1}+\frac{16x+31}{x^2+3x+4}=\frac{7(5-2x)}{(x^2+x+1)(x^2+3x+4)}$$
Similarly
$$\frac{9+5x}{x^2+2x+3}-\frac{14+5x}{x^2+3x+4}=\frac{2(2x-3)}{(x^2+2x+3)(x^2+3x+4)}$$

whence from equation (iv) it follows that the required partial fractions are

$$-\frac{1}{7}\frac{16x-1}{x^2+x+1}+\frac{1}{2}\frac{5x+9}{x^2+2x+3}-\frac{3}{14}\frac{x+12}{x^2+3x+4}.$$

Example (c). To find the partial fractions of $\dfrac{1}{x^6-1}$.

We have

$$x^6-1 = (x^2-1)(x^4+x^2+1) = (x-1)(x+1)((x^2+1)^2-x^2)$$
$$= (x-1)(x+1)(x^2+x+1)(x^2-x+1).$$

The partial fractions corresponding to the factors $x-1$ and $x+1$ are $\dfrac{1}{6.1^5}\dfrac{1}{x-1}$ and $\dfrac{1}{6(-1)^5}\dfrac{1}{x+1}$, i.e. $\dfrac{1}{6}\dfrac{1}{x-1}$ and $-\dfrac{1}{6}\dfrac{1}{x+1}$.

Now

$$\frac{1}{x^6-1}-\left\{\frac{1}{6(x-1)}-\frac{1}{6(x+1)}\right\} = \frac{1}{x^6-1}-\frac{1}{3(x^2-1)}$$
$$= -\frac{x^4+x^2-2}{3(x^6-1)} = -\frac{x^2+2}{3(x^4+x^2+1)}.$$

Furthermore

$$\frac{1}{x^2-x+1}+\frac{1}{x^2+x+1} = \frac{2(x^2+1)}{x^4+x^2+1},$$
$$\frac{1}{x^2-x+1}-\frac{1}{x^2+x+1} = \frac{2x}{x^4+x^2+1},$$

therefore

$$\frac{x-2}{x^2-x+1}-\frac{x+2}{x^2+x+1} = \frac{2x^2-4(x^2+1)}{x^4+x^2+1} = -\frac{2(x^2+2)}{x^4+x^2+1}.$$

Thus the required expression in partial fractions is

$$\frac{1}{6(x-1)}-\frac{1}{6(x+1)}+\frac{1}{6}\frac{x-2}{x^2-x+1}-\frac{1}{6}\frac{x+2}{x^2+x+1}.$$

If we need only the partial fraction corresponding to one particular factor of the denominator we may proceed as follows. Suppose that x^2+px+q is a factor of $Q(x)$ and that

$$Q(x) = (x^2+px+q)Q^*(x).$$

If $\dfrac{Ax+B}{x^2+px+q}$ is the required partial fraction then

$$\frac{P(x)}{Q(x)}-\frac{Ax+B}{x^2+px+q} = \frac{P(x)-(Ax+B)Q^*(x)}{Q(x)}$$

must be expressible in the form $\dfrac{P^*(x)}{Q^*(x)}$ and therefore
$$P(x)-(Ax+B)Q^*(x)$$
is divisible by x^2+px+q. Let $lx+m$ and $rx+s$ be the remainders when $P(x)$ and $Q^*(x)$ are divided by x^2+px+q, then we require that $lx+m-(Ax+B)(rx+s)$ be divisible by x^2+px+q. Since
$$lx+m-(Ax+B)(rx+s)$$
$$= -Ar(x^2+\{(As+Br-l)/Ar\}x+(Bs-m)/Ar)$$
we must have $Bs-m = -Aqr$ and $As+Br-l = -Apr$, whence A and B are determined.

For instance, to find the partial fraction of $\dfrac{1}{x^6-1}$ corresponding to the factor x^2+x+1 we require to find A and B so that
$$(Ax+B)(x^2-1)(x^2-x+1)-1$$
is divisible by x^2+x+1. Now
$$(x^2-1)(x^2-x+1) = \{x^2+x+1-(x+2)\}\{x^2+x+1-2x\}$$
and $2x(x+2) = 2(x^2+x+1+x-1)$ and so we require that
$$2(x-1)(Ax+B)-1$$
be divisible by x^2+x+1, i.e. that $2Ax^2+2(B-A)x-(2B+1)$ is so divisible, and therefore $B-A = A$, $2B+1 = -2A$, whence
$$A = -\tfrac{1}{6}, \qquad B = -\tfrac{1}{3}$$
and the partial fraction is $-\dfrac{1}{6}\dfrac{x+2}{x^2+x+1}$.

As a final illustration we shall obtain the partial fractions of
$$1/\{x^{2n}-2x^n\cos\alpha+1\}.$$

Starting from the formula†

$$x^{2n}-2x^n\cos\alpha+1$$
$$=\left(x^2-2x\cos\frac{\alpha}{n}+1\right)\left(x^2-2x\cos\frac{\alpha+2\pi}{n}+1\right)\left(x^2-2x\cos\frac{\alpha+4\pi}{n}+1\right)\ldots$$
$$\ldots\left(x^2-2x\cos\frac{\alpha+2(n-1)\pi}{n}+1\right),$$
we have
$$\log(x^{2n}-2x^n\cos\alpha+1) = \sum_{r=1}^{n}\log\left(x^2-2x\cos\frac{\alpha+2r\pi}{n}+1\right), \qquad \text{(i)}$$

† See Example 5.21.

and differentiating with respect to x, we find

$$\frac{nx^{n-1}(x^n-\cos\alpha)}{x^{2n}-2x^n\cos\alpha+1} = \sum_1^n \frac{x-\cos\{(\alpha+2r\pi)/n\}}{x^2-2x\cos\{(\alpha+2r\pi)/n\}+1},$$

and therefore multiplying by x and subtracting both sides from n

$$\frac{x^n\cos\alpha-1}{x^{2n}-2x^n\cos\alpha+1} = \frac{1}{n}\sum_1^n \frac{x\cos\{(\alpha+2r\pi)/n\}-1}{x^2-2x\cos\{(\alpha+2r\pi)/n\}+1}.$$

Differentiating (i) with respect to α,

$$\frac{x^n\sin\alpha}{x^{2n}-2x^n\cos\alpha+1} = \frac{1}{n}\sum_1^n \frac{x\sin\{(\alpha+2r\pi)/n\}}{x^2-2x\cos\{(\alpha+2r\pi)/n\}+1}$$

and so

$$\frac{1}{x^{2n}-2x^n\cos\alpha+1} = \frac{x^n\sin\alpha\cos\alpha}{\sin\alpha(x^{2n}-2x^n\cos\alpha+1)} - \frac{x^n\cos\alpha-1}{x^{2n}-2x^n\cos\alpha+1}$$

$$= \frac{1}{n\sin\alpha}\sum_1^n \frac{\sin\alpha+x\sin[\{2r\pi-(n-1)\alpha\}/n]}{x^2-2x\cos\{(\alpha+2r\pi)/n\}+1}.$$

Observe that the method of Example (*b*) still applies if two of the quadratic factors in the denominator are equal, showing how the case of repeated trinomial factors may be treated.

In both Examples (*a*) and (*b*) we considered, for the sake of symmetry, both the sum and difference of the quadratic factors Q_1 and Q_2 (say), but it suffices to consider only the difference $1/Q_1-1/Q_2$ in conjunction with either one of the two identities $1/Q_1 = Q_2/Q_1Q_2$, $1/Q_2 = Q_1/Q_1Q_2$.

VII

SUCCESSIVE DIFFERENTIATION

MAXIMA AND MINIMA

7. It may happen that the derivative $f'(x)$ of some function $f(x)$ is itself a differentiable function with a derivative which we shall denote by $f''(x)$; $f''(x)$ is called the second derivative of $f(x)$. If $f''(x)$ is also differentiable, then its derivative $f'''(x)$ is called the third derivative of $f(x)$, and so on. Thus if each function of the sequence $f(x), f'(x), f''(x), \ldots, f^n(x), \ldots$ is the derivative of its predecessor, then $f^n(x)$ is the nth derivative of $f(x)$. We shall also denote the nth derivative of a function $f(x)$ by $D_x^n f(x)$ or by $\dfrac{d^n}{dx^n} f(x)$.

For example, since e^x is its own derivative, the nth derivative of e^x is e^x itself; and since each term of the sequence

$$1/x, \quad -1/x^2, \quad 2!/x^3, \quad -3!/x^4, \quad \ldots, \quad (-1)^n n!/x^{n+1}, \quad \ldots$$

is the derivative of its predecessor, therefore $(-1)^n n!/x^{n+1}$ is the nth derivative of $1/x$.

7.1. If p is *not* a positive integer, the nth derivative of x^p is $p(p-1)(p-2)\ldots(p-n+1)x^{p-n}$ for any value of n, but if p *is* a positive integer the nth derivative of x^p is

$$p(p-1)(p-2)\ldots(p-n+1)x^{p-n}, \quad p!, \quad \text{or} \quad 0$$

according as n is less than, equal to, or greater than p. Accordingly the nth derivative of a polynomial is zero if the degree of the polynomial is less than n.

7.11. If we consider in turn the successive derivatives of $\sin x$ we obtain the sequence $\cos x$, $-\sin x$, $-\cos x$, $\sin x$,... so that the nth derivative of $\sin x$ is $\cos x$, $-\sin x$, $-\cos x$, or $\sin x$ according as n leaves the remainder 1, 2, 3, or 0 when divided by 4; since $\sin(x+n\pi/2)$ takes precisely these values for these values of n we have the simple formula

$$D^n \sin x = \sin(x+n\pi/2).$$

Similarly

$$D^n \cos x = \cos(x+n\pi/2).$$

7.2. If the nth derivatives $u_n(x)$ and $v_n(x)$ of two functions $u(x)$ and $v(x)$ are known then the nth derivative of the product function $u(x)v(x)$ is given by the *Leibnitz formula*:

7.21. $$D^n uv = u_n v + \binom{n}{1} u_{n-1} v_1 + \binom{n}{2} u_{n-2} v_2 + \\ + \binom{n}{3} u_{n-3} v_3 + \ldots + \binom{n}{1} u_1 v_{n-1} + uv_n.$$

For $n = 1$ this gives $D(uv) = u_1 v + uv_1$, which agrees with the formula we have previously proved for the derivative of a product.

If the formula is true for $n = k$ then

$$D^{k+1}uv = D(D^k uv)$$

$$= D\left\{u_k v + \binom{k}{1} u_{k-1} v + \ldots + \binom{k}{r-1} u_{k-r+1} v_{r-1} + \binom{k}{r} u_{k-r} v_r + \ldots + uv_k\right\}$$

$$= u_{k+1} v + \left\{1 + \binom{k}{1}\right\} u_k v_1 + \ldots + \left\{\binom{k}{r-1} + \binom{k}{r}\right\} u_{k+1-r} v_r + \ldots + uv_{k+1}$$

$$= u_{k+1} v + \binom{k+1}{1} u_k v_1 + \ldots + \binom{k+1}{r} u_{k+1-r} v_r + \ldots + uv_{k+1}$$

since

$$\binom{k}{r-1} + \binom{k}{r} = k!/(r-1)!(k-r+1)! + k!/r!(k-r)!$$

$$= k!\{r+k-r+1\}/r!(k-r+1)! = (k+1)!/r!(k+1-r)! = \binom{k+1}{r},$$

and therefore 7.21 is true for $n = k+1$, whence it follows that it is true for any value of n.

EXAMPLE. The nth derivative of e^x/x is

$$e^x/x - \binom{n}{1} e^x/x^2 + \binom{n}{2} 2!\, e^x/x^3 - \ldots + (-1)^n n!\, e^x/x^{n+1}$$

$$= \{x^n - nx^{n-1} + n(n-1)x^{n-2} - \\ - n(n-1)(n-2)x^{n-3} + \ldots + (-1)^n n!\} e^x/x^{n+1}.$$

7.3. There is no simple formula expressing the nth derivative of $f(\phi(x))$ in terms of the derivatives of $f(x)$ and $\phi(x)$; the following result, however, is useful in particular cases.

7.31. $D_x^n f(u) = \sum_{r=1}^{n} P_r(u) f^r(u)$, where u stands for $\phi(x)$ and $P_r(u)$ is *independent* of the function $f(x)$.

For $n = 1$, 7.31 is true with $P_1(u) = u'$; if it is true for $n = k$ then

$$D_x^{k+1}f(u) = D_x(D_x^k f(u)) = D_x(P_1 f'(u)+P_2 f''(u)+\ldots+P_k f^k(u))$$
$$= u'P_1' f'+u'(P_1+P_2')f''+u'(P_2+P_3')f'''+\ldots+$$
$$+u'(P_{k-1}+P_k')f^k+u'P_k f^{k+1},$$

and so since $(P_{r-1}+P_r')u'$ is independent of $f(x)$ if P_{r-1} and P_r are independent of $f(x)$, the formula is true for $n = k+1$, and therefore true for all values of n.

7.311.* The coefficients $P_r(u)$ depend not only upon r and u but also upon n; since they do not depend upon $f(x)$ they may be determined by giving special values to $f(x)$. For instance, take $f(x) = x$ and we have

$$D_x^n u = P_1(u). \tag{i}$$

Next take $f(x) = x^2$ and we find

$$D_x^n u^2 = P_1 2u+P_2.2; \tag{ii}$$

and then $f(x) = x^3$, giving

$$D_x^n u^3 = P_1 3u^2+P_2 6u+P_3 6; \tag{iii}$$

and so from (i), (ii), (iii) in turn we find

$$P_1 = D_x^n u,$$
$$P_2 = \frac{1}{2!}D_x^n u^2-uD_x^n u,$$
$$P_3 = \frac{1}{3!}D_x^n u^3-\frac{u}{2!}D_x^n u^2+\frac{u^2}{2!}D_x^n u,$$

and so on. The values of $D_x^n u^2, D_x^n u^3,\ldots$ are given in turn by the Leibnitz formula, considering u^2 as $u.u$, u^3 as $u.u^2$, and so on.

Example. To find the nth derivative of $f(\log x)$.

Write u for $\log x$. By formula 7.31

$$D_x^n f(\log x) = P_1 D_u f(u)+P_2 D_u^2 f(u)+\ldots+P_n D_u^n f(u),$$

where P_r does not depend upon $f(x)$.

Take $f(x) = e^{ax}$ so that $f(\log x) = x^a$ and $f(u) = e^{au}$.

Since $D_x^n x^a = a(a-1)(a-2)\ldots(a-n+1)x^{a-n}$ and

$$D_u^r e^{au} = a^r e^{au} = a^r x^a$$

we have

$$(aP_1+a^2P_2+\ldots+a^nP_n)x^a = a(a-1)\ldots(a-n+1)x^{a-n},$$

i.e. $$aP_1+a^2P_2+\ldots+a^nP_n = a(a-1)\ldots(a-n+1)/x^n,$$

so that P_r is the coefficient of a^r in the expansion of

$$a(a-1)\ldots(a-n+1)/x^n$$

and therefore P_r is the coefficient of $D_u^r f(u)$ in

$$\frac{1}{x^n}D_u(D_u-1)(D_u-2)\ldots(D_u-n+1)f(u),$$

i.e.

$$(P_1 D_u+P_2 D_u^2+P_3 D_u^3+\ldots+P_n D_u^n)f(u)$$
$$= \frac{1}{x^n}D_u(D_u-1)\ldots(D_u-n+1)f(u),$$

which proves that the nth derivative of $f(\log x)$ is

$$\frac{1}{x^n}D_u(D_u-1)\ldots(D_u-n+1)f(u).$$

7.4. If $f(x+h) = a_0(x)+ha_1(x)+\frac{h^2}{2!}a_2(x)+\frac{h^3}{3!}a_3(x)+\ldots+\frac{h^n}{n!}a_n(x)+\ldots$ the series being convergent for $h = h_0$ (and any x in an interval i) then $f^n(x)$, the nth derivative of $f(x)$, is equal to $a_n(x)$.

For $\frac{d}{dh}f(x+h) = f'(x+h)\frac{d}{dh}(x+h) = f'(x+h)$ and therefore

$$f^n(x+h) = \frac{d^n}{dh^n}f(x+h).$$

But by Theorem 3.7

$$\frac{d}{dh}f(x+h) = a_1(x)+ha_2(x)+\frac{h^2}{2!}a_3(x)+\ldots+\frac{h^{n-1}}{(n-1)!}a_n(x)+\ldots$$

for $|h| < |h_0|$, and therefore

$$\frac{d^n}{dh^n}f(x+h) = a_n(x)+ha_{n+1}(x)+\frac{h^2}{2!}a_{n+2}(x)+\ldots$$

for $|h| < |h_0|$, i.e. $f^n(x+h) = a_n(x)+ha_{n+1}(x)+\ldots$, in particular $f^n(x) = a_n(x)$, and $f(x) = a_0(x)$. Furthermore, taking $x = 0$, we have

$$f(h) = f(0)+hf'(0)+\frac{h^2}{2!}f''(0)+\frac{h^3}{3!}f'''(0)+\ldots+\frac{h^n}{n!}f^n(0)+\ldots$$

(provided $x = 0$ lies in the interval i).

It is important to notice that in proving this theorem we have assumed that $f(x+h)$ is given by a series of powers of h. The

question as to what functions $f(x+h)$ have such an expansion is a much more difficult one, and is considered in a later chapter on Taylor's theorem.

EXAMPLE. To find the nth derivative of e^{ax^2}.

$$e^{a(x+h)^2}-e^{ax^2} = e^{ax^2}\{e^{ah(2x+h)}-1\}$$

$$= e^{ax^2}\left[a\frac{h}{1!}(2x+h)+a^2\frac{h^2}{2!}(2x+h)^2+\ldots+a^n\frac{h^n}{n!}(2x+h)^n+\ldots\right],$$

the series being absolutely convergent for all h and x.

The coefficient of $\frac{h^n}{n!}$, and therefore the nth derivative of e^{ax^2}, is

$$e^{ax^2}\left[(2x)^n a^n+\frac{n(n-1)}{1}(2x)^{n-2}a^{n-1}+\right.$$

$$\left.+\frac{n(n-1)(n-2)(n-3)}{1.2}(2x)^{n-4}a^{n-2}+\ldots\right].$$

By means of formula 7.31 we can now deduce the nth derivative of $f(x^2)$.

For $D^n f(x^2) = \sum_{k=1}^{n} P_k(u)f^k(u)$, where $u = x^2$. Since $P_k(u)$ is independent of f we may take $f(u) = e^{au}$, in which case $P_k(u)$ is the coefficient of $a^k e^{au}$ in $D^n e^{ax^2}$, and therefore

$$D^n f(x^2) = (2x)^n f^n(u)+\frac{n(n-1)}{1}(2x)^{n-2}f^{n-1}(u)+$$

$$+\frac{n(n-1)(n-2)(n-3)}{1.2}(2x)^{n-4}f^{n-2}(u)+\ldots$$

7.5. The nth derivatives of the functions $\frac{1}{1+x^2}$, $\frac{x}{1+x^2}$ are of considerable importance and may be expressed in particularly simple form. We shall show:

7.501. $$D_x^n\{1/(1+x^2)\} = (-1)^n n!\sin^{n+1}\theta\sin(n+1)\theta,$$
$$\theta = \arcsin(1+x^2)^{-\frac{1}{2}}.$$

7.502. $$D_x^n\{x/(1+x^2)\} = (-1)^n n!\sin^{n+1}\theta\cos(n+1)\theta,$$
$$\theta = \arcsin(1+x^2)^{-\frac{1}{2}}.$$

Proof of 7.501. $D\{1/(1+x^2)\} = -2x/(1+x^2)^2$.

$$(-1)1!\sin^2\theta\sin 2\theta = -2\sin^3\theta\cos\theta = -2x/(1+x^2)^2$$

since $\sin\theta = 1/(1+x^2)^{\frac{1}{2}}$

Thus 7.501 is true for $n = 1$; if it is true for $n = p$, then

$$\begin{aligned} D_x^{p+1}\{1/(1+x^2)\} &= D\{D_x^p(1/(1+x^2))\} \\ &= \frac{d}{d\theta}\{(-1)^p p!\sin^{p+1}\theta\sin(p+1)\theta\}\frac{d\theta}{dx} \\ &= (-1)^p(p+1)!\{\sin^p\theta\cos\theta\sin(p+1)\theta+ \\ &\qquad +\sin^{p+1}\theta\cos(p+1)\theta\}\{-1/(1+x^2)\} \\ &= (-1)^{p+1}(p+1)!\sin^p\theta\sin(p+2)\theta\sin^2\theta \\ &= (-1)^{p+1}(p+1)!\sin^{p+2}\theta\sin(p+2)\theta, \end{aligned}$$

so that 7.501 is true for $n = p+1$ and therefore true for all values of n.

7.502 is proved in a similar way.

The connexion of the derivative of $1/(1+x^2)$ with trigonometric functions is a consequence of the derivative of $\arctan x$ being $1/(1+x^2)$.

7.51. By means of the formulae (Examples V)

$$\sin(n+1)\theta = \binom{n+1}{1}\cos^n\theta\sin\theta-\binom{n+1}{3}\cos^{n-2}\theta\sin^3\theta+\ldots,$$

$$\cos(n+1)\theta = \cos^{n+1}\theta-\binom{n+1}{2}\cos^{n-1}\theta\sin^2\theta+\binom{n+1}{4}\cos^{n-3}\theta\sin^4\theta-\ldots$$

we can express 7.501 and 7.502 in the form:

7.511. $$D_x^n\{1/(1+x^2)\} = (-1)^n n!\times$$
$$\times\left\{\binom{n+1}{1}x^n-\binom{n+1}{3}x^{n-2}+\binom{n+1}{5}x^{n-4}-\ldots\right\}\Big/(1+x^2)^{n+1}.$$

7.512. $$D_x^n\{x/(1+x^2)\} = (-1)^n n!\times$$
$$\times\left\{x^{n+1}-\binom{n+1}{2}x^{n-1}+\binom{n+1}{4}x^{n-3}-\ldots\right\}\Big/(1+x^2)^{n+1}.$$

In the same way we can show, for $a > 0$:

7.513. $$D_x^n\{1/(a+x^2)\} = (-1)^n n!\times$$
$$\times\left\{\binom{n+1}{1}x^n-\binom{n+1}{3}x^{n-2}a+\binom{n+1}{5}x^{n-4}a^2-\binom{n+1}{7}x^{n-6}a^3+\ldots\right\}\div$$
$$\div(a+x^2)^{n+1}.$$

7.514. $$D_x^n\{x/(a+x^2)\} = (-1)^n n!\times$$
$$\times\left\{x^{n+1}-\binom{n+1}{2}x^{n-1}a+\binom{n+1}{4}x^{n-3}a^2-\binom{n+1}{6}x^{n-5}a^3+\ldots\right\}\div$$
$$\div(a+x^2)^{n+1}.$$

7.52. Since $D_a^p\{1/(a+x^2)\} = (-1)^p p!/(a+x^2)^{p+1}$, if we *assume* that $D_x^n D_a^p\{1/(a+x^2)\} = D_a^p D_x^n\{1/(a+x^2)\}$ then the value of

$$D_x^n\{1/(a+x^2)^{p+1}\}$$

may be obtained by differentiating the right-hand side of 7.513 p times with respect to a. Each term is of the form $a^r/(a+x^2)^{n+1}$ and the kth derivative, with respect to a, of both a^r and $1/(a+x^2)^{n+1}$ may be written down immediately, so that the pth derivative of $a^r/(a+x^2)^{n+1}$ is given by the Leibnitz formula.

7.6. We have seen in the previous chapter how to express any rational function $P(x)/Q(x)$ as a sum of terms each of which is either a polynomial or of the form $A/(x+a)^p$ or of the form

$$(Ax+B)/\{(x+\lambda)^2+a\}^p.$$

In §§ 7 and 7.1 we showed how to find the nth derivatives of a polynomial and of $A/(x+a)^p$ and in §§ 7.5 and 7.52 we showed how to find the nth derivatives of $1/(x^2+a)^p$ and of $x/(x^2+a)^p$; since the derivatives with respect to x of $(x+\lambda)/\{(x+\lambda)^2+a\}^p$ and

$$1/\{(x+\lambda)^2+a\}^p$$

are the same as the derivatives with respect to t of $t/(t^2+a)^p$ and $1/(t^2+a)^p$, where $t = x+\lambda$, it follows that the nth derivative of any rational function may be obtained by expressing the rational function in its partial fractions.

7.7. Maxima and minima

The maxima of a function $f(x)$ are the values of $f(x)$ which are greatest in their immediate neighbourhood and the minima are those which are least in their immediate neighbourhood. Formally we define:

$f(a)$ is a maximum value of the function $f(x)$ if $f(a) > f(x)$ for *all* values of x sufficiently near a and $f(a)$ is a minimum if $f(a) < f(x)$ for *all* values of x sufficiently near a.

A maximum value of a function is not necessarily greater than a minimum, for a maximum is the greatest value only in relation to a small enough region.

For example, consider the function $f(x) = 3x^5-25x^3+60x$. A simple calculation shows that

$$\text{(i)}\quad f(1)-f(x) = (x-1)^2\{45-[5(x-1)+15(x-1)^2+3(x-1)^3]\}$$
$$> 22(x-1)^2 \quad \text{if } |x-1| < 1$$
$$> 0 \quad \text{provided } x \neq 1.$$

(ii) $\quad f(x)-f(2) = (x-2)^2\{90+[95(x-2)+30(x-2)^2+3(x-2)^3]\}$
$$> 34(x-2)^2 \quad \text{if } |x-2| < \tfrac{1}{2}$$
$$> 0 \quad \text{provided } x \neq 2.$$

(iii) $\quad f(x)-f(-1) = (x+1)^2\{45+[5(x+1)-15(x+1)^2+3(x+1)^3]\}$
$$> 22(x+1)^2 \quad \text{if } |x+1| < 1$$
$$> 0 \quad \text{provided } x \neq -1.$$

(iv) $\quad f(-2)-f(x) = (x+2)^2\{90-[95(x+2)-30(x+2)^2+3(x+2)^3]\}$
$$> 34(x+2)^2 \quad \text{if } |x+2| < \tfrac{1}{2}$$
$$> 0 \quad \text{provided } x \neq -2.$$

From (i) and (iv) it follows that $f(x)$ has a maximum at $x = 1$ and at $x = -2$ and from (ii) and (iii) that $f(x)$ has a minimum at $x = 2$ and at $x = -1$. Since $f(2) = 16$ and $f(-2) = -16$ we see that the minimum value of $f(x)$ at $x = 2$ is *greater* than the maximum value at $x = -2$.

7.71. If $f(x)$ is a differentiable function and if $f(a)$ is either a maximum or a minimum value of $f(x)$ then $f'(a) = 0$.

Suppose that $f(a)$ is a maximum; then $f(a) > f(a \pm 1/n)$. But $\dfrac{f(a\pm 1/n)-f(a)}{\pm 1/n} \to f'(a)$ and so since $\dfrac{f(a+1/n)-f(a)}{1/n} < 0$, $f'(a) \leqslant 0$, and since $\dfrac{f(a-1/n)-f(a)}{-1/n} > 0$, $f'(a) \geqslant 0$, whence it follows that $f'(a) = 0$.

Similarly, if $f(a)$ is a minimum, $f'(a) = 0$.

Theorem 7.71 may be expressed by saying that the values of x for which $f(x)$ is a maximum or a minimum are amongst the roots of the equation $f'(x) = 0$.

7.711. Although $f'(a) = 0$ is a *necessary* condition for $f(a)$ to be a maximum or minimum value of a differentiable function $f(x)$ it is not a *sufficient* condition. For example the derivative of x^3 is $3x^2$ which vanishes for $x = 0$, but the origin is neither a maximum nor a minimum value of x^3 for $x^3 > 0$ if $x > 0$ and $x^3 < 0$ if $x < 0$ however small $|x|$ may be.

7.72. If $f'(a) = 0$, $f'(a+h) < 0$, and $f'(a-h) > 0$ for all small positive values of h then $f(a)$ is a *maximum* value of $f(x)$.

Proof. From $f'(a) = 0$, $f'(a+h) < 0$ for $h \leqslant h_0$ it follows that $f(x)$ is steadily decreasing in the interval $(a, a+h_0)$ and therefore $f(a) > f(x)$ in this interval. From $f'(a) = 0$, $f'(a-h) > 0$ for $h \leqslant h_0$ it follows that $f(x)$ is steadily increasing in the interval $(a-h_0, a)$ and therefore $f(x) < f(a)$ in this interval. Thus $f(a) > f(x)$ throughout $(a-h_0, a+h_0)$, $x \neq a$, and so $f(a)$ is a maximum.

A similar argument shows that:

7.721. If $f'(a) = 0$, $f'(a+h) > 0$, and $f'(a-h) < 0$ for all small positive values of h then $f(a)$ is a minimum value of $f(x)$.

7.722. If $f'(a) = 0$ but $f'(x)$ has the same sign on either side of $x = a$ then $f(x)$ has neither a maximum nor minimum at $x = a$.

Suppose for instance that $f'(x)$ is positive on either side of $x = a$. Then $f(x)$ is increasing throughout some interval $(a-h_0, a+h_0)$ and so $f(a-h) < f(a) < f(a+h)$, which proves that $f(x)-f(a)$ is positive or negative according as x is greater than or less than a and so $f(a)$ is neither a maximum nor minimum value of $f(x)$.

7.73. If $f'(a) = 0$ then $f(a)$ is a minimum or maximum value of $f(x)$ according as $f''(a) > 0$ or $f''(a) < 0$.

For if $f''(a) > 0$ then $f'(x)$ is *increasing* at the point $x = a$, so that $f'(a+h) > f'(a) = 0$ and $f'(a-h) < f'(a) = 0$ for all small positive values of h, and if $f''(a) < 0$ then $f'(x)$ is *decreasing* at the point $x = a$, so that $f'(a+h) < f'(a) = 0$ and $f'(a-h) > f'(a) = 0$.

7.74. If $f'(a) = f''(a) = 0$, and $f''(a)$ is a maximum value of $f''(x)$, then $f(a)$ is a maximum value of $f(x)$; and if $f''(a)$ is a minimum value then $f(a)$ is a minimum value.

For if $f''(a) = 0$ is a maximum value of $f''(x)$, then $f''(x) < 0$ for all x near a, so that $f'(x)$ decreases near a. Since $f'(a) = 0$, therefore when $x < a, f'(x) > 0$, and when $x > a, f'(x) < 0$, whence by 7.72, $f(a)$ is a maximum value of $f(x)$. Similarly, if $f''(a)$ is a minimum then $f(a)$ is a minimum.

7.741. If $f'(a) = 0$ and $f'(a)$ is a maximum or minimum value of $f'(x)$ then $f(a)$ is neither a maximum nor minimum value of $f(x)$.

For if $f'(a) = 0$ is a maximum or a minimum, then $f'(x)$ has the same sign on either side of the point $x = a$, and therefore, by 7.722, $f(x)$ has neither a maximum nor minimum value at $x = a$.

7.75. If $f'(a) = f''(a) = \dots = f^{2n-1}(a) = 0$, and $f^{2n}(a) \neq 0$, then $f(a)$ is maximum or minimum according as $f^{2n}(a)$ is negative or positive.

If $f^{2n}(a) > 0$, then since $f^{2n-1}(a) = 0$, it follows, by 7.73, that $f^{2n-2}(a)$ is a minimum value of $f^{2n-2}(x)$. Hence, by 7.74, $f^{2n-4}(a)$ is a minimum, and therefore by repeated application of 7.74, $f^{2n-6}(a)$, $f^{2n-8}(a)$,..., $f''(a)$, $f(a)$ are all minimum values. Thus if $f^{2n}(a) > 0$, $f(a)$ is a minimum value of $f(x)$. Similarly, if $f^{2n}(a) < 0$ then $f(a)$ is a maximum value.

7.751. If $f'(a) = f''(a) = \dots = f^{2n}(a) = 0$, and $f^{2n+1}(a) \neq 0$, then $f(a)$ is neither a maximum nor minimum value of $f(x)$.

Suppose that $f^{2n+1}(a) < 0$; then $f^{2n-1}(a)$ is a maximum value of $f^{2n-1}(x)$, by 7.73, and therefore, by repeated application of 7.74, $f'(a)$ is a maximum value of $f'(x)$ and so, by 7.741, $f(a)$ is neither a maximum nor minimum value of $f(x)$. Similarly, if $f^{2n+1}(a) > 0$, then $f'(a)$ is a minimum and $f(a)$ is neither a maximum nor minimum value of $f(x)$.

7.76. If $f''(a) = 0$ and $f''(x)$ changes sign as x passes through the value a, then $x = a$ is called a *point of inflexion* of $f(x)$.

7.761. If $f''(a) = f'''(a) = \dots = f^{2n}(a) = 0$, and $f^{2n+1}(a) \neq 0$, then $x = a$ is a point of inflexion of $f(x)$.

For, by 7.74, $f'''(a)$ is a maximum or minimum value of $f'''(x)$; hence $f''(x)$ is steadily increasing or steadily decreasing near $x = a$, so that $f''(x)$ changes sign as x passes through the value a.

Examples. To find the maximum and minimum values of $x(a-b)/(x-a)(x-b)$, $a > b > 0$.

$$f(x) = \frac{(a-b)x}{(x-a)(x-b)} = \frac{a}{x-a} - \frac{b}{x-b},$$

$$f'(x) = -\frac{a}{(x-a)^2} + \frac{b}{(x-b)^2},$$

$$f'' = \frac{2a}{(x-a)^3} - \frac{2b}{(x-b)^3}.$$

The roots of $f'(x) = 0$ are $x = \sqrt{(ab)}$ and $x = -\sqrt{(ab)}$:

$$f''\{\sqrt{(ab)}\}$$

$$= \frac{2}{\sqrt{a}(\sqrt{b}-\sqrt{a})^3} - \frac{2}{\sqrt{b}(\sqrt{a}-\sqrt{b})^3} = -\frac{2}{(\sqrt{a}-\sqrt{b})^3}\left\{\frac{1}{\sqrt{a}} + \frac{1}{\sqrt{b}}\right\} < 0,$$

$$f''\{-\sqrt{(ab)}\}$$

$$= -\frac{}{\sqrt{a}(\sqrt{a}+\sqrt{b})^3}+\frac{}{\sqrt{b}(\sqrt{a}+\sqrt{b})^3} = \frac{2}{(\sqrt{a}+\sqrt{b})^3}\frac{\sqrt{a}-\sqrt{b}}{\sqrt{(ab)}} > 0.$$

Thus $(a-b)x/(x-a)(x-b)$ has a maximum for $x = \sqrt{(ab)}$ and a minimum for $x = -\sqrt{(ab)}$ and the maximum and minimum values are

$$-\frac{\sqrt{a}+\sqrt{b}}{\sqrt{a}-\sqrt{b}} \quad \text{and} \quad -\frac{\sqrt{a}-\sqrt{b}}{\sqrt{a}+\sqrt{b}}, \quad \text{of which the latter is the greater.}$$

To find the maximum and minimum values of $(x-a)^3(x-b)^4$, when $a > b > 0$.

$$f(x) = (x-a)^3(x-b)^4,$$

$$f'(x) = 3(x-a)^2(x-b)^4+4(x-a)^3(x-b)^3$$

$$= (x-a)^2(x-b)^3\{7x-4a-3b\},$$

$$f''(x) = 6(x-a)(x-b)^2\{2(x-a)^2+4(x-a)(x-b)+(x-b)^2\}.$$

The roots of $f'(x) = 0$ are $x = a$, $x = b$, $x = (4a+3b)/7$.

Since $x-a = -\frac{3}{7}(a-b)$ and $x-b = \frac{4}{7}(a-b)$, when

$$x = (4a+3b)/7,$$

therefore $f''\{(4a+3b)/7\} = 2^6.3^2.(a-b)^5/7^4 > 0$ so that $f(x)$ has a *minimum* at $x = (4a+3b)/7$.

Since $f''(a) = 0$ we must examine the sign of $f'(x)$ on either side of $x = a$. As $(4a+3b)/7$ lies between a and b, therefore when x is near a, $x-(4a+3b)/7$ is positive, and so too $(x-b)^3$ is positive; since $(x-a)^2$ is also positive when x is near to but different from a, therefore $f'(x)$ is positive on either side of $x = a$ which proves that $f(x)$ has neither a maximum nor minimum value at $x = a$.

Near b, $x-(4a+3b)/7$ is negative and $(x-b)^3$ is positive or negative according as x is greater or less than b; therefore $f'(b-h)$ is positive, $f'(b+h)$ is negative, for small enough positive values of h, which proves that $f(x)$ has a *maximum* value at $x = b$.

Alternative proof. By Leibnitz's theorem

$$f'''(x) = \{D^3(x-a)^3\}(x-b)^4+3D^2(x-a)^3D(x-b)^4+$$
$$+3D(x-a)^3D^2(x-b)^4+(x-a)^3D^3(x-b)^4,$$

so that $\qquad f'''(a) = 6(a-b)^4 \quad \text{and} \quad f'''(b) = 0,$

and

$$f^{\text{iv}}(x) = \{D^4(x-b)^4\}(x-a)^3+4D^3(x-b)^4D(x-a)^3+$$
$$+6D^2(x-b)^4D^2(x-a)^3+4D(x-b)^4D^3(x-a)^3,$$

so that $$f^{\text{iv}}(b) = 24(b-a)^3.$$

Thus $$f'(a) = f''(a) = 0 \quad \text{and} \quad f'''(a) = 6(a-b)^4,$$

and so, by 7.751, $x = a$ is a point of inflexion of $f(x)$.

Since

$$f'(b) = f''(b) = f'''(b) = 0, \quad \text{and} \quad f^{\text{iv}}(b) = 24(b-a)^3 < 0,$$

therefore, by 7.75, $f(b)$ is a maximum value of $f(x)$.

VIII

THE INDEFINITE INTEGRAL

INTEGRATION BY PARTS. CHANGE OF VARIABLE. REDUCTION FORMULAE. RATIONAL FUNCTIONS. QUADRATIC IRRATIONALS

8. If $f(x)$ is the derivative of $F(x)$ then $F(x)$ is said to be an *integral* of $f(x)$. There are several notations for an integral, corresponding to the notations D, D_x, d/dx for derivatives, viz. I, ${}^xI$, $\int (\ )\, dx$. The origin of I or ${}^xI$ as the integral sign is self-evident, but the origin of the sign $\int (\ )\, dx$ will not become apparent until we have approached the question of integration from an entirely different point of view, though its technical utility is readily grasped.

It follows from Theorem 3.651 that if $F(x)$ and $G(x)$ are two integrals of a function then $F(x)$ and $G(x)$ are equal, apart from an additive constant, for $F'(x) = G'(x) = f(x)$, say. Accordingly, if one integral of a function is known, every integral is known. In writing down an integral we shall generally ignore an additive constant; thus, for instance, we shall write

$$I2x = x^2$$

rather than $I2x = x^2+5$ or $I2x = x^2+a$, which are of course equally true, though on the other hand we shall usually find it more convenient to write $I2(x+1) = (x+1)^2$ rather than

$$I2(x+1) = x^2+2x.$$

8.01. Differentiation and integration are *inverse* operations in the same sense in which we spoke of *inverse* functions, for, by definition,

$$I\{Df(x)\} = f(x) \quad \text{and} \quad D\{If(x)\} = f(x)$$

whatever differentiable function $f(x)$ may be.

8.02. Instead of $\int 1\, dx$, i.e. ${}^xI\, 1$, we shall write $\int dx$, and instead of $\int \{1/f(x)\}\, dx$ we shall (sometimes) write $\int dx/f(x)$.

8.1. Since the derivatives of $\dfrac{x^{n+1}}{n+1}$, $\sin x$, $-\cos x$, $\tan x$, $-\cot x$, $\tan^{-1}x$, $-\cot^{-1}x$, $\sin^{-1}x$, $-\cos^{-1}x$, e^x, $\log x$ are x^n, $\cos x$, $\sin x$, $\sec^2 x$, $\operatorname{cosec}^2 x$, $1/(1+x^2)$, $1/(1+x^2)$, $1/\sqrt{(1-x^2)}$, $1/\sqrt{(1-x^2)}$, e^x, $1/x$,

respectively, it follows that the integrals of the latter functions are the former, and so (provided $n \neq -1$)

$$\int x^n\,dx = \frac{x^{n+1}}{n+1}, \quad \int \cos x\,dx = \sin x, \quad \int \sin x\,dx = -\cos x;$$

$$\int \sec^2 x\,dx = \tan x, \qquad \int \operatorname{cosec}^2 x\,dx = -\cot x,$$

$$\int \frac{1}{1+x^2}\,dx = \tan^{-1} x,$$
$$= -\cot^{-1} x \quad \text{(apart from an additive constant)};$$

$$\int dx/\sqrt{(1-x^2)} = \sin^{-1} x$$
$$= -\cos^{-1} x \quad \text{(apart from an additive constant)};$$

$$\int e^x\,dx = e^x, \qquad \int (1/x)\,dx = \log x, \quad \text{provided } x > 0.$$

8.2. $$\int \{f(x)+g(x)\}\,dx = \int f(x)\,dx + \int g(x)\,dx.$$

For $$\frac{d}{dx}\left[\int \{f(x)+g(x)\}\,dx\right] = f(x)+g(x)$$

and

$$\frac{d}{dx}\left[\int f(x)\,dx + \int g(x)\,dx\right] = \frac{d}{dx}\int f(x)\,dx + \frac{d}{dx}\int g(x)\,dx$$
$$= f(x)+g(x)$$

and so both sides of equation 8.2 are integrals of the same function, and so are equal, apart from an additive constant.

8.21. It follows from 8.2 that

$$\int \{f(x)+g(x)+h(x)\}\,dx = \int f(x)\,dx + \int g(x)\,dx + \int h(x)\,dx,$$

for

$$\int \{f(x)+g(x)+h(x)\}\,dx = \int \{f(x)+g(x)\}\,dx + \int h(x)\,dx$$
$$= \int f(x)\,dx + \int g(x)\,dx + \int h(x)\,dx.$$

Similarly, we can show that $\int \left\{\sum_1^n u_r(x)\right\} dx = \sum_1^n \int u_r(x)\,dx$ for any n.

EXAMPLES.

$$\int (p+qx+rx^2+sx^3)\,dx = px+\frac{qx^2}{2}+\frac{rx^3}{3}+\frac{sx^4}{4}.$$

$$\int dx/\sin^2 x \cos^2 x$$
$$= \int [\{\sin^2 x + \cos^2 x\}/\sin^2 x \cos^2 x]\, dx = \int \left\{\frac{1}{\sin^2 x} + \frac{1}{\cos^2 x}\right\} dx$$
$$= \int \operatorname{cosec}^2 x\, dx + \int \sec^2 x\, dx = \tan x - \cot x.$$

Change of variable. Integration by substitution.

If $\phi(t)$ has a derivative $\phi'(t)$ and if $x = \phi(t)$ then:

8.3. $$\int f(x)\, dx = \int f(\phi(t))\, \phi'(t)\, dt.$$

For

$$\frac{d}{dt}\left\{\int f(x)\, dx\right\} = \frac{d}{dx}\left\{\int f(x)\, dx\right\} \frac{dx}{dt} = f(x)\frac{dx}{dt}$$
$$= f(\phi(t))\phi'(t), \quad \text{since } x = \phi(t),$$

and $$\frac{d}{dt}\int \{f(\phi(t))\phi'(t)\}\, dt = f(\phi(t))\phi'(t),$$

and therefore $\int f(x)\, dx$ and $\int f(\phi(t))\phi'(t)\, dt$ are equal apart from an additive constant.

EXAMPLES. (*a*) Take $x = \dfrac{at}{b}$, then

$$\int \frac{dx}{a^2+b^2x^2} = \int \frac{1}{a^2+a^2t^2}\frac{dx}{dt}\, dt = \frac{1}{ab}\int \frac{1}{1+t^2}\, dt$$
$$= \frac{1}{ab}\tan^{-1} t = \frac{1}{ab}\tan^{-1}\frac{bx}{a}.$$

(*b*) Take $x = (t-p)^2+q^2$, then

$$\int \frac{dx}{x} = \int \frac{1}{(t-p)^2+q^2}\frac{dx}{dt}\, dt = \int \frac{2(t-p)\, dt}{(t-p)^2+q^2}$$

and therefore

$$\int \frac{2(t-p)\, dt}{(t-p)^2+q^2} = \log x = \log\{(t-p)^2+q^2\}.$$

(*c*) Take $x = f(t) > 0$, then

$$\int \frac{dx}{x} \quad \int \frac{1}{f(t)}\frac{dx}{dt}\, dt = \int \frac{f'(t)}{f(t)}\, dt.$$

Thus $$\int \frac{f'(t)}{f(t)}\, dt = \log x = \log f(t).$$

(*d*) To evaluate $\int \sqrt{(a^2-x^2)}\,dx$. Take $x = a\sin t$, then

$$\int \sqrt{(a^2-x^2)}\,dx = \int \sqrt{(a^2-a^2\sin^2 t)}\,\frac{dx}{dt}\,dt = a^2\int \cos^2 t\,dt$$

$$= \frac{a^2}{2}\int (1+\cos 2t)\,dt = \frac{a^2}{2}\left(t+\frac{\sin 2t}{2}\right) \quad \text{(taking } a>0\text{)}.$$

Since $\sin t = x/a$, if t lies in the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$ then

$$t = \sin^{-1}\frac{x}{a}$$

and $\cos t = \sqrt{(1-\sin^2 t)}$ and so

$$t+\frac{\sin 2t}{2} = t+\sin t\cos t = \sin^{-1}\frac{x}{a}+\frac{x}{a}\sqrt{\left(1-\frac{x^2}{a^2}\right)}.$$

Thus
$$\int \sqrt{(a^2-x^2)}\,dx = \frac{a^2}{2}\sin^{-1}\frac{x}{a}+\frac{x}{2}\sqrt{(a^2-x^2)}.$$

(*e*) $\int \frac{1}{a+bt}\,dt$. Take $a+bt = x$, then

$$\int \frac{1}{x}\,dx = \int \frac{1}{a+bt}\,\frac{dx}{dt}\,dt = \int \frac{b}{a+bt}\,dt$$

and so

$$\int \frac{1}{a+bt}\,dt = \frac{1}{b}\log x = \frac{1}{b}\log(a+bt), \quad \text{provided } a+bt>0.$$

Similarly

$$\int \frac{1}{a-bt}\,dt = -\frac{1}{b}\log(a-bt), \quad \text{provided } a-bt>0.$$

Hence

$$\int \frac{2a}{a^2-b^2t^2}\,dt = \int \frac{1}{a-bt}\,dt + \int \frac{1}{a+bt}\,dt = \frac{1}{b}\log\frac{a+bt}{a-bt},$$

provided $a^2 > b^2t^2$,

and

$$\int \frac{2bt}{a^2-b^2t^2}\,dt = \int \frac{1}{a-bt}\,dt - \int \frac{1}{a+bt}\,dt = -\frac{1}{b}\log(a^2-b^2t^2),$$

provided $a^2 > b^2t^2$.

Note that $$\int \frac{1}{t-a}\,dt = \log(t-a) \quad \textit{if} \quad t>a,$$

but $$\int \frac{1}{t-a}\,dt = \int \frac{-1}{a-t}\,dt = \log(a-t), \quad \textit{if} \quad t<a,$$

since the logarithm is defined only for a positive argument.

Integration by parts:

8.4. $$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx.$$

Proof. Since

$$\frac{d}{dx}(f(x)g(x)) = f'(x)g(x)+f(x)g'(x)$$

therefore $$f(x)g(x) = \int \{f'(x)g(x)+f(x)g'(x)\}\,dx$$

$$= \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx,$$

i.e. $$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx,$$

which is known as the formula for *integration by parts.*

EXAMPLES.

(*a*) $$\int xe^x\,dx = \int \left(\frac{d}{dx}e^x\right)x\,dx = xe^x - \int e^x\,dx = xe^x - e^x,$$

i.e. $$\int xe^x\,dx = e^x(x-1).$$

(*b*) $$\int x^n e^{ax}\,dx = x^n\frac{e^{ax}}{a} - \frac{n}{a}\int x^{n-1}e^{ax}\,dx$$

$$= x^n\frac{e^{ax}}{a} - \frac{n}{a^2}x^{n-1}e^{ax} + \frac{n(n-1)}{a^2}\int x^{n-2}e^{ax}\,dx,$$

and so on, giving

$$\int x^n e^{ax}\,dx = \frac{e^{ax}}{a}\Bigg\{x^n - n\frac{x^{n-1}}{a} +$$

$$+n(n-1)\frac{x^{n-2}}{a^2} - \ldots(-1)^r n(n-1)\ldots(n-r+1)\frac{x^{n-r}}{a^r}\ldots+(-1)^n\frac{n!}{a^n}\Bigg\}.$$

(*c*) $$\int x\sin x\,dx = -x\cos x + \int \cos x\,dx = \sin x - x\cos x.$$

(*d*) $$\int e^{ax}\sin bx\,dx = \frac{e^{ax}}{a}\sin bx - \frac{b}{a}\int e^{ax}\cos bx\,dx$$

$$= \frac{e^{ax}}{a}\sin bx - \frac{b}{a^2}e^{ax}\cos bx - \frac{b^2}{a^2}\int e^{ax}\sin bx\,dx$$

and therefore

$$\left(1+\frac{b^2}{a^2}\right)\int e^{ax}\sin bx\,dx = \frac{e^{ax}}{a^2}(a\sin bx - b\cos bx),$$

$$\int e^{ax}\sin bx\,dx = \frac{e^{ax}}{a^2+b^2}(a\sin bx - b\cos bx).$$

(e) $$\int \log x\,dx = \int (Dx)\log x\,dx = x\log x - \int x.\frac{1}{x}\,dx$$
$$= x(\log x - 1).$$

(f) $$\int \tan^{-1}x\,dx = \int (Dx)\tan^{-1}x\,dx = x\tan^{-1}x - \int \frac{x}{1+x^2}\,dx$$
$$= x\tan^{-1}x - \tfrac{1}{2}\log(1+x^2).$$

8.5. Integration of the circular functions

The integrals of $\tan x$ and $\cot x$.

$$\int \tan x\,dx = \int \frac{\sin x}{\cos x}\,dx = -\int \frac{(d/dx)\cos x}{\cos x}\,dx = -\log\cos x,$$

in any interval in which $\cos x > 0$, i.e.

$$\int \tan x\,dx = -\log\cos x.$$

Similarly $$\int \cot x\,dx = +\log\sin x.$$

The integrals of $\operatorname{cosec} x$ and $\sec x$ are slightly more difficult.

$$\int \frac{1}{\sin x}\,dx = \frac{1}{2}\int \frac{1}{\sin\frac{1}{2}x\cos\frac{1}{2}x}\,dx = \frac{1}{2}\int \frac{\sec^2\frac{1}{2}x}{\tan\frac{1}{2}x}\,dx = \int \frac{D\tan\frac{1}{2}x}{\tan\frac{1}{2}x}\,dx$$
$$= \log\tan\tfrac{1}{2}x, \quad \text{in any interval in which } \tan\tfrac{1}{2}x > 0,$$

$$\int \operatorname{cosec} x\,dx = \log\tan\tfrac{1}{2}x.$$

Writing $x = \frac{1}{2}\pi + t$ we have

$$\log\tan(\tfrac{1}{4}\pi+\tfrac{1}{2}t) = \int \frac{1}{\sin(t+\frac{1}{2}\pi)}\frac{dx}{dt}\,dt = \int \frac{1}{\cos t}\,dt,$$

i.e.

$$\int \sec t\,dt = \log\tan(\tfrac{1}{4}\pi+\tfrac{1}{2}t),$$ in any interval in which $\tan(\frac{1}{4}\pi+\frac{1}{2}t)$ is positive.

Since

$$\tan(\tfrac{1}{4}\pi+\tfrac{1}{2}t) = \frac{1+\tan\frac{1}{2}t}{1-\tan\frac{1}{2}t} = \frac{\cos\frac{1}{2}t+\sin\frac{1}{2}t}{\cos\frac{1}{2}t-\sin\frac{1}{2}t} = \frac{(\cos\frac{1}{2}t+\sin\frac{1}{2}t)^2}{\cos^2\frac{1}{2}t-\sin^2\frac{1}{2}t}$$
$$= \frac{1+\sin t}{\cos t} = \sec t+\tan t,$$

therefore $$\int \sec t\,dt = \log(\sec t+\tan t).$$

8.51. The integral of $1/(a+b\cos x)$.

Write $t = \tan\frac{1}{2}x$ so that

$$\cos x = (1-t^2)/(1+t^2) \quad \text{and} \quad \frac{dt}{dx} = \tfrac{1}{2}\sec^2\tfrac{1}{2}x = \tfrac{1}{2}(1+t^2)$$

and we have

$$\int \frac{1}{a+b\cos x}\,dx = \int \frac{1}{a+b\{(1-t^2)/(1+t^2)\}}\,\frac{2}{1+t^2}\,dt$$

$$= 2\int \frac{dt}{a+b+t^2(a-b)} = \frac{2}{a-b}\int \frac{dt}{t^2+(a+b)/(a-b)},$$

provided $a \neq b$.

The value of the integral depends upon the sign of

$$\frac{a+b}{a-b} = \frac{(a+b)^2}{a^2-b^2},$$

and so upon the sign of a^2-b^2.

If $a^2 > b^2$, let $\dfrac{a+b}{a-b} = \lambda^2 > 0$, then, since $\displaystyle\int \frac{\lambda}{t^2+\lambda^2}\,dt = \tan^{-1}\frac{t}{\lambda}$, we have

$$\int \frac{1}{a+b\cos x}\,dx = \frac{2}{\lambda(a-b)}\tan^{-1}\frac{t}{\lambda} = \frac{2}{\sqrt{(a^2-b^2)}}\tan^{-1}\left\{t\sqrt{\left(\frac{a-b}{a+b}\right)}\right\}$$

where $t = \tan\frac{1}{2}x$.

If $a^2 < b^2$, let $\dfrac{a+b}{a-b} \quad -\mu^2 < 0$, then, since

$$\int \frac{1}{t^2-\mu^2}\,dt = \frac{1}{2\mu}\log\frac{t-\mu}{t+\mu} \quad \text{or} \quad \frac{1}{2\mu}\log\frac{\mu-t}{\mu+t},$$

according as $t > \mu$ or $t < \mu$, we have

$$\int \frac{1}{a+b\cos x}\,dx = \frac{1}{\mu(a-b)}\log\frac{t-\mu}{t+\mu} \quad \text{or} \quad \frac{}{\mu(a-b)}\log\frac{\mu-t}{\mu+t}$$

$$= \frac{1}{\sqrt{(b^2-a^2)}}\log\left[\frac{t+\sqrt{\{(b+a)/(b-a)\}}}{t-\sqrt{\{(b+a)/(b-a)\}}}\right]$$

$$\text{or} \quad \frac{1}{\sqrt{(b^2-a^2)}}\log\left[\frac{\sqrt{\{(b+a)/(b-a)\}}+t}{\sqrt{\{(b+a)/(b-a)\}}-t}\right]$$

according as $t > \sqrt{\{(b+a)/(b-a)\}}$ or $t < \sqrt{\{(b+a)/(b-a)\}}$.

These results may be transformed in various ways. For instance, if we write

$$t = \tan\tfrac{1}{2}x = \sqrt{\left(\frac{a+b}{a-b}\right)}\tan u,$$

then
$$\frac{a\cos x+b}{a+b\cos x}=\frac{(a+b)-(a-b)t^2}{(a+b)+(a-b)t^2}=\frac{1-\tan^2 u}{1+\tan^2 u}=\cos 2u$$
and therefore
$$2\tan^{-1}\left\{\sqrt{\left(\frac{a-b}{a+b}\right)}\,t\right\}=2u=\cos^{-1}\left\{\frac{a\cos x+b}{a+b\cos x}\right\},$$
whence it follows that
$$\int\frac{dx}{a+b\cos x}=\frac{1}{\sqrt{(a^2-b^2)}}\cos^{-1}\left\{\frac{a\cos x+b}{a+b\cos x}\right\},$$
provided $a^2>b^2$, i.e. $|a|>|b|$.

If $a=b$, the value of the integral is $(1/a)\int dt=(1/a)\tan\frac{1}{2}x$, $a\neq 0$, and if $a=-b$ the value is $(1/a)\int(1/t^2)\,dt=-(1/a)\cot\frac{1}{2}x$, $a\neq 0$.

The integration of $1/\{a+b\cos x+c\sin x\}$ may be reduced to the preceding case by writing $\theta=\tan^{-1}(c/b)$, $r=+\sqrt{(b^2+c^2)}$, so that
$$b\cos x+c\sin x=r(\cos\theta\cos x+\sin\theta\sin x)=r\cos(x-\theta),$$
and therefore if $y=x-\theta$,
$$\int\frac{1}{a+b\cos x+c\sin x}\,dx=\int\frac{dy}{a+r\cos y}.$$

8.52. The integral of $\sin^m x\cos^n x$.

8.521. If m is an odd positive integer $=2k+1$, say, write $t=\cos x$, then
$$\int\sin^m x\cos^n x\,dx=\int(\sin^2 x)^k\cos^n x\sin x\,dx=-\int(1-t^2)^k t^n\,dt,$$
which is integrated by expanding $(1-t^2)^k$.

8.522. If $n=2k+1$, write $t=\sin x$, then
$$\int\sin^m x\cos^n x\,dx=\int t^m(1-t^2)^k\,dt.$$

8.523. If $m+n=-2k$, where k is a positive integer, write $t=\tan x$, then
$$\int\sin^m x\cos^n x\,dx=\int\left(\frac{\sin x}{\cos x}\right)^m\cos^{m+n}x\,dx$$
$$=\int\tan^m x(\sec^2 x)^{k-1}\sec^2 x\,dx=\int t^m(1+t^2)^{k-1}\,dt,$$
which is integrated by expanding $(1+t^2)^{k-1}$.

8.524. If $m+n = 0$, $m > 0$, and integral, write $t = \tan x$, then

$$\int \sin^m x \cos^n x \, dx = \int \frac{t^m}{1+t^2} \, dt.$$

Divide t^m by $1+t^2$, giving a quotient $t^{m-2}-t^{m-4}+t^{m-6}-\ldots$ and leaving a remainder ± 1 or $\pm t$ according as m is even or odd, whence the integral is evaluated.

8.525. If $m+n = 0$, $n > 0$, and integral, write $u = \cot x$, then

$$\int \sin^m x \cos^n x \, dx = -\int \frac{u^n}{1+u^2} \, du,$$

which is evaluated as in 8.524.

8.526. If $m = -(2k+1)$, $n = 0$, write $t = \tan \frac{1}{2}x$, then

$$\int \frac{dx}{\sin^{2k+1} x} = \frac{1}{2^{2k}} \int \frac{(1+t^2)^{2k}}{t^{2k+1}} \, dt,$$

which is integrated by expanding $(1+t^2)^{2k}$.

8.527. If $m = 0$, $n = -(2k+1)$, write $y = \frac{1}{2}\pi - x$, then

$$\int \frac{dx}{\cos^{2k+1} x} = -\int \frac{dy}{\sin^{2k+1} y}.$$

8.528. If m and n are both integers, positive or negative (or zero) the integral of $\sin^m x \cos^n x$ is evaluated by means of the following *reduction formulae.*

Since $\int \sin^m x \cos^n x \, dx = \int \sin^{m-1} x \cos^n x \sin x \, dx$, integrating by parts we have, provided $n \neq -1$,

(1) $\int \sin^m x \cos^n x \, dx$

$$\frac{\sin^{m-1} x \cos^{n+1} x}{n+1} + \frac{m-1}{n+1} \int \sin^{m-2} x \cos^{n+2} x \, dx.$$

But

$$\int \sin^{m-2} x \cos^{n+2} x \, dx = \int \sin^{m-2} x \cos^n x (1-\sin^2 x) \, dx$$

$$= \int \sin^{m-2} x \cos^n x \, dx - \int \sin^m x \cos^n x \, dx,$$

and therefore

$$\frac{m+n}{n+1} \int \sin^m x \cos^n x \, dx$$

$$= -\frac{\sin^{m-1} x \cos^{n+1} x}{n+1} + \frac{m-1}{n+1} \int \sin^{m-2} x \cos^n x \, dx,$$

whence, provided $m+n \neq 0$,

(2) $$\int \sin^m x \cos^n x \, dx = -\frac{\sin^{m-1}x \cos^{n+1}x}{m+n} + \frac{m-1}{m+n} \int \sin^{m-2}x \cos^n x \, dx.$$

In the same way we can show that, provided $m \neq -1$,

(3) $$\int \sin^m x \cos^n x \, dx = \frac{\sin^{m+1}x \cos^{n-1}x}{m+1} + \frac{n-1}{m+1} \int \sin^{m+2}x \cos^{n-2}x \, dx,$$

and, provided $m+n \neq 0$,

(4) $$\int \sin^m x \cos^n x \, dx = \frac{\sin^{m+1}x \cos^{n-1}x}{m+n} + \frac{n-1}{m+n} \int \sin^m x \cos^{n-2}x \, dx$$

Formulae (2) and (4) may also be written in the forms

(5) $$\int \sin^{m-2}x \cos^n x \, dx = \frac{\sin^{m-1}x \cos^{n+1}x}{m-1} + \frac{m+n}{m-1} \int \sin^m x \cos^n x \, dx, \quad \text{provided } m \neq 1,$$

(6) $$\int \sin^m x \cos^{n-2}x \, dx = -\frac{\sin^{m+1}x \cos^{n-1}x}{n-1} + \frac{m+n}{n-1} \int \sin^m x \cos^n x \, dx, \quad \text{provided } n \neq 1.$$

If m is a positive integer the repeated application of formula (2) reduces the exponent of $\sin x$ to zero or unity without changing the exponent of $\cos x$; thus the integral of $\sin^m x \cos^n x$ is in this case reduced to one of the integrals

$$\int \sin x \cos^n x \, dx = -\frac{\cos^{n+1}x}{n+1} \quad \text{and} \quad \int \cos^n x \, dx.$$

If n is positive then $\int \cos^n x \, dx$ is reduced to either $\int 1 \, dx = x$ or to $\int \cos x = \sin x$ by repeated application of (4).

If m is a negative integer the repeated application of formula (5) raises the exponent of $\sin x$ to zero or -1, and so the integral is reduced either to $\int \cos^n x \, dx$ or to

$$\int \frac{\cos^n x}{\sin x} \, dx = \int \frac{\cos^n x}{\sin^2 x} \sin x \, dx = -\int \frac{t^n}{1-t^2} \, dt,$$

writing $t = \cos x$, and if n is positive this is evaluated by dividing

$1-t^2$ into t^n. If n is negative, $\int \cos^n x\,dx$ is reduced by formula (6) to one of the integrals

$$\int \sec^2 x\,dx = \tan x \quad \text{or} \quad \int \frac{1}{\cos x}\,dx = \log(\sec x+\tan x)$$

and $\int \frac{\cos^n x}{\sin x}\,dx$ is reduced either to

$$\int \frac{1}{\sin x\cos x}\,dx = \log\tan x \quad \text{or to} \quad \int \frac{1}{\sin x}\,dx = \log\tan \tfrac{1}{2}x.$$

Thus the integral is evaluated when m and n are any integers whatsoever.

If m and n are of opposite sign the reduction process may be more speedily effected by a preliminary use of either formula (1) or (3).

8.6. Miscellaneous integrals

The integral of $x(\tan x)^2$.

$$\int x\tan^2 x\,dx = \int x\sec^2 x\,dx - \int x\,dx = x\tan x - \int \tan x\,dx - \tfrac{1}{2}x^2$$
$$= x\tan x + \log\cos x - \tfrac{1}{2}x^2.$$

The integral of $1/(a+b\tan x)$.

$$V = \int \frac{dx}{a+b\tan x} = \int \frac{\cos x\,dx}{a\cos x+b\sin x}$$
$$= \frac{1}{a}\int \frac{a\cos x+b\sin x}{a\cos x+b\sin x}\,dx - \frac{b}{a}\int \frac{\sin x\,dx}{a\cos x+b\sin x}$$
$$= \frac{x}{a}-\frac{b}{a}\int \frac{\sin x\,dx}{a\cos x+b\sin x}.$$

Therefore

$$\log(a\cos x+b\sin x) = \int \frac{b\cos x-a\sin x}{a\cos x+b\sin x}\,dx$$
$$= bV+\frac{a^2}{b}\left(V-\frac{x}{a}\right) = (a^2+b^2)V/b - ax/b,$$

whence $$V = \frac{b}{a^2+b^2}\log(a\cos x+b\sin x)+\frac{ax}{a^2+b^2}.$$

The integral of $x^2/(x\cos x-\sin x)^2$.

Since $$\frac{d}{dx}\left(\frac{1}{x\cos x-\sin x}\right) = \frac{x\sin x}{(x\cos x-\sin x)^2}$$

therefore

$$\int \frac{x^2}{(x\cos x-\sin x)^2}\,dx$$
$$= \int \frac{x}{\sin x}\frac{d}{dx}\left(\frac{1}{x\cos x-\sin x}\right)dx$$
$$= \frac{x}{\sin x(x\cos x-\sin x)} - \int \frac{d}{dx}\left(\frac{x}{\sin x}\right)\frac{1}{x\cos x-\sin x}\,dx$$
$$= \frac{x}{\sin x(x\cos x-\sin x)} + \int \frac{dx}{\sin^2 x} = \frac{x-\cos x(x\cos x-\sin x)}{\sin x(x\cos x-\sin x)}$$
$$= \frac{x\sin x+\cos x}{x\cos x-\sin x} = \frac{\tan x+1/x}{1-(\tan x)/x} = \tan\{x+\tan^{-1}(1/x)\}.$$

The integral of $(1-r\cos x)/(1-2r\cos x+r^2)$

$$\int \frac{1-r\cos x}{1-2r\cos x+r^2}\,dx = \frac{1}{2}\int dx + \frac{1}{2}\int \frac{1-r^2}{1-2r\cos x+r^2}\,dx.$$

Writing $t = \tan\frac{1}{2}x$ we have

$$\frac{1}{2}\int \frac{1-r^2}{1-2r\cos x+r^2}\,dx = \int \frac{1-r^2}{(1-r)^2+(1+r)^2t^2}\,dt = \tan^{-1}\frac{1+r}{1-r}t.$$

Therefore

$$\int \frac{1-r\cos x}{1-2r\cos x+r^2}\,dx = \frac{x}{2}+\tan^{-1}\left\{\frac{1+r}{1-r}\tan\tfrac{1}{2}x\right\}.$$

8.7. The integral of a rational function

By expressing a rational function $P(x)/Q(x)$ in *partial fractions* the integral of $P(x)/Q(x)$ is transformed into a sum of integrals of the forms $\int x^p\,dx$, $\int \frac{dx}{(x+a)^p}$, and $\int \frac{x+c}{\{(x+\lambda)^2+a^2\}^p}\,dx$; accordingly the integral of any rational function may be determined provided only these integrals can be evaluated.† The integrals $\int x^p\,dx$, $\int \frac{1}{(x+a)^{p+1}}\,dx$, $p>0$, and $\int \frac{1}{x+a}\,dx$ are trivial, their values being $\frac{x^{p+1}}{p+1}$, $-1/p(x+a)^p$, and $\log(x+a)$ respectively; so too the integral

$$\int \frac{x+c}{\{(x+\lambda)^2+a^2\}}\,dx = \frac{1}{2}\int \frac{2(x+\lambda)}{(x+\lambda)^2+a^2}\,dx + \frac{c-\lambda}{a}\int \frac{a}{(x+\lambda)^2+a^2}\,dx$$
$$= \tfrac{1}{2}\log\{(x+\lambda)^2+a^2\} + \frac{c-\lambda}{a}\tan^{-1}\frac{x+\lambda}{a}.$$

† Unless the resolution into partial fractions is impracticable, which may happen either because the factors of $Q(x)$ are too numerous and diverse or because the equation $Q(x) = 0$ is not soluble except by approximations to its roots.

There remains to consider the integral of $(x+c)/\{(x+\lambda)^2+a^2\}^p$, $p > 1$. To evaluate this integral we express $x+c$ in the form

$$A\{(x+\lambda)^2+a^2\}-2(x+\lambda)\tfrac{1}{2}(Ax-B),$$

where $\qquad A = (c-\lambda)/a^2, \qquad B = (\lambda^2-\lambda c+a^2)/a^2,$

so that

$$\int \frac{x+c}{\{(x+\lambda)^2+a^2\}^p}\,dx$$

$$= A\int \frac{1}{\{(x+\lambda)^2+a^2\}^{p-1}}\,dx - \frac{1}{2}\int \frac{2(x+\lambda)}{\{(x+\lambda)^2+a^2\}^p}(Ax-B)\,dx$$

$$= A\int \frac{}{\{(x+\lambda)^2+a^2\}^{p-1}}\,dx + \frac{1}{2(p-1)}\frac{Ax-B}{\{(x+\lambda)^2+a^2\}^{p-1}}$$

$$-\frac{A}{2(p-1)}\int \frac{1}{\{(x+\lambda)^2+a^2\}^{p-1}}\,dx, \quad \text{after integrating by parts,}$$

$$= \frac{1}{2(p-1)}\frac{Ax-B}{\{(x+\lambda)^2+a^2\}^{p-1}} + \frac{(2p-3)}{2(p-1)}A\int \frac{1}{\{(x+\lambda)^2+a^2\}^{p-1}}\,dx.$$

Repeating the process the exponent of the denominator is step by step reduced to unity.

EXAMPLE. To evaluate

$$\int \frac{x^2+3x+4}{(x^2+1)^3}\,dx$$

We have $\qquad x^2+3x+4 = (x^2+1)+3(x+1),$

and $\qquad x+1 = x^2+1-2x\,.\,\tfrac{1}{2}(x-1)$

whence

$$\int \frac{x^2+3x+4}{(x^2+1)^3}\,dx = \int \frac{1}{(x^2+1)^2}\,dx + 3\int \frac{(x^2+1)-2x\,.\,\frac{1}{2}(x-1)}{(x^2+1)^3}\,dx$$

$$= 4\int \frac{1}{(x^2+1)^2}\,dx - \frac{3}{2}\int \frac{2x}{(x^2+1)^3}(x-1)\,dx$$

$$= 4\int \frac{1}{(x^2+1)^2}\,dx + \frac{3}{4}\frac{x-1}{(x^2+1)^2} - \frac{3}{4}\int \frac{1}{(x^2+1)^2}\,dx$$

$$= \frac{3}{4}\frac{x-1}{(x^2+1)^2} + \frac{13}{4}\int \frac{1}{(x^2+1)^2}\,dx.$$

Since
$$1 = (x^2+1)-2x.\tfrac{1}{2}x,$$
$$\int \frac{1}{(x^2+1)^2}\,dx = \int \frac{1}{x^2+1}\,dx - \frac{1}{2}\int \frac{2x}{(x^2+1)^2}x\,dx$$
$$= \int \frac{1}{x^2+1}\,dx + \frac{1}{2}\frac{x}{(x^2+1)} - \frac{1}{2}\int \frac{1}{x^2+1}\,dx$$
$$= \frac{1}{2}\frac{x}{x^2+1} + \tfrac{1}{2}\tan^{-1}x.$$
Hence
$$\int \frac{x^2+3x+4}{(x^2+1)^3}\,dx = \frac{3}{4}\frac{x-1}{(x^2+1)^2} + \frac{13}{8}\frac{x}{x^2+1} + \frac{13}{8}\tan^{-1}x.$$

The same method of reduction may be applied to any integral of the form $\int \frac{P}{Q^r}\,dx$, where P and Q are polynomials and Q does not contain a repeated factor. We have but to choose polynomials C and D so that $P = CQ+DQ'$, where Q' is the derivative of Q, and then
$$\int \frac{P}{Q^r}\,dx = \int \frac{C}{Q^{r-1}}\,dx + \int \frac{Q'}{Q^r}D\,dx$$
$$= \int \frac{C}{Q^{r-1}}\,dx - \frac{1}{r-1}\frac{D}{Q^{r-1}} + \frac{1}{r-1}\int \frac{D'}{Q^{r-1}}\,dx$$
$$= -\frac{1}{r-1}\frac{D}{Q^{r-1}} + \int \left(C+\frac{D'}{r-1}\right)\frac{1}{Q^{r-1}}\,dx.$$

Of course the integral of $(x+c)/\{(x+\lambda)^2+a^2\}^p$ may also be evaluated by the substitution $x+\lambda = a\tan z$, but this method is generally far more laborious than the former.

8.71. If $R(x, y)$ is a rational function then the integral of
$$R(\sin x, \cos x)$$
may be evaluated by the substitution $t = \tan\frac{1}{2}x$, for the integral becomes
$$\int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2\,dt}{1+t^2},$$
which is of the type discussed in § 8.7.

8.8. Integration of quadratic irrationals

8.81. We consider first the integrals of the three functions
$$1/\sqrt{(a^2-x^2)}, \qquad 1/\sqrt{(a^2+x^2)}, \qquad 1/\sqrt{(x^2-a^2)}, \qquad a > 0.$$

By the substitution $x = a\sin t$,

$$\int \frac{1}{\sqrt{(a^2-x^2)}}\,dx = \int \frac{1}{\sqrt{(1-\sin^2 t)}}\cos t\,dt;$$

if we suppose that t lies in the interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, then

$$\cos t = \sqrt{(1-\sin^2 t)} \quad \text{and} \quad t = \sin^{-1}\frac{x}{a},$$

whence
$$\int \frac{1}{\sqrt{(a^2-x^2)}}\,dx = \int dt = t = \sin^{-1}\frac{x}{a}.$$

The substitution $x = a\,\mathrm{sh}\,t$ gives

$$\int \frac{1}{\sqrt{(a^2+x^2)}}\,dx = \int \frac{1}{\mathrm{ch}\,t}\mathrm{ch}\,t\,dt = t = \mathrm{sh}^{-1}\frac{x}{a} = \log\{x+\sqrt{(a^2+x^2)}\}$$

If $x = a\,\mathrm{ch}\,t$, $t \geqslant 0$, then

$$\mathrm{sh}\,t = \sqrt{(\mathrm{ch}^2 t-1)} \quad \text{and} \quad t = \mathrm{ch}^{-1}\frac{x}{a} = \log\{x+\sqrt{(x^2-a^2)}\}$$

and therefore

$$\int \frac{1}{\sqrt{(x^2-a^2)}}\,dx = \int dt = t = \log\{x+\sqrt{(x^2-a^2)}\}.$$

8.811. The integral of $1/\sqrt{(ax^2+2bx+c)}$.

(i) If $a > 0$,

$$\sqrt{(ax^2+2bx+c)} = \sqrt{a}\sqrt{\left\{\left(x+\frac{b}{a}\right)^2+\frac{ac-b^2}{a^2}\right\}};$$

we must distinguish the cases $ac < b^2$, $ac = b^2$, $ac > b^2$.

If $ac < b^2$ let $ac-b^2 = -\alpha^2 a^2$, and so

$$\int \frac{1}{\sqrt{(ax^2+2bx+c)}}\,dx = \frac{1}{\sqrt{a}}\log\left[x+\frac{b}{a}+\sqrt{\left\{\left(x+\frac{b}{a}\right)^2-\alpha^2\right\}}\right]$$

$$= \frac{1}{\sqrt{a}}\log\left\{x+\frac{b}{a}+\sqrt{(ax^2+2bx+c)}/\sqrt{a}\right\}$$

$$= \frac{1}{\sqrt{a}}\log[ax+b+\sqrt{\{a(ax^2+2bx+c)\}}]$$

apart from an additive constant.

If $ac > b^2$ the result is unchanged.

(ii) If $a < 0$ write $a = -\lambda^2$, then

$$ax^2+2bx+c = c+b^2/\lambda^2-(\lambda x-b/\lambda)^2 = c+b^2/\lambda^2-\lambda^2(x-b/\lambda^2)^2$$

and therefore

$$\int \frac{1}{\sqrt{(ax^2+2bx+c)}}\,dx = \frac{1}{\lambda}\sin^{-1}\frac{\lambda(x-b/\lambda^2)}{\sqrt{(c+b^2/\lambda^2)}} \quad \text{where } \lambda = \sqrt{(-a)}.$$

If $ac = b^2$, $a > 0$, then

$$\sqrt{(ax^2+2bx+c)} = \sqrt{a}(x+b/a) \quad \text{if } x > -b/a,$$

and therefore

$$\int \frac{1}{\sqrt{(ax^2+2bx+c)}}\,dx = \frac{1}{\sqrt{a}}\log(x+b/a).$$

8.82. Writing $y = \sqrt{(ax^2+2bx+c)}$ we have

$$\int \frac{px+q}{y}\,dx = \frac{p}{2a}\int \frac{2(ax+b)}{y}\,dx + (q-pb/a)\int \frac{1}{y}\,dx$$

$$= \frac{p}{a}(ax^2+2bx+c)^{\frac{1}{2}} + (q-bp/a)\int \frac{1}{y}\,dx,$$

whence the solution follows from 8.811.

8.83.
$$a\int y\,dx = (ax+b)y - \int \frac{(ax+b)^2}{y}\,dx$$

$$= (ax+b)y - a\int y\,dx + (ac-b^2)\int \frac{dx}{y}$$

and the solution follows from 8 811 and 8.82.

8.84. A reduction formula for $\int \frac{x^m}{y}\,dx$.

Write $I_m = \int \frac{x^m}{y}\,dx$, then since

$$y\frac{d}{dx}(x^{m-1}y) = max^m + (2m-1)bx^{m-1} + c(m-1)x^{m-2},$$

we have $$I_m = \frac{x^{m-1}}{am}y - b\frac{(2m-1)}{am}I_{m-1} - \frac{c(m-1)}{am}I_{m-2}.$$

The values of I_0 and I_1 are given by 8.811 and 8.82.

8.85. The integral of $1/(x-p)^m y$.

Write $x-p = 1/t$, then $\log(x-p) = -\log t$ and so

$$\frac{1}{x-p}\frac{dx}{dt} = -\frac{1}{t},$$

whence

$$\int \frac{1}{(x-v)^m u}\,dx = -\int \frac{t^{m-1}}{v}\frac{1}{t}\,dt$$

$$= -\int \frac{t^{m-1}}{\sqrt{\{(ap^2+2bp+c)t^2+2(ap+b)t+a\}}}\,dt,$$

which is of the form 8.84 if $m > 1$, and of the form 8.811 if $m = 1$.

8.86. We consider next the integral of $(px+q)/(ax^2+b)\sqrt{(x^2+c)}$, which we evaluate in two parts.

First we evaluate $\displaystyle\int \frac{x\,dx}{(ax^2+b)\sqrt{(x^2+c)}}$.

Write $\sqrt{(x^2+c)} = y$ so that $ax^2+b = ay^2+b-ac$ and $x\dfrac{dx}{dy} = y$ and the integral becomes

$$\int \frac{dy}{ay^2+b-ac} = \frac{1}{a}\int \frac{dy}{y^2+k}, \quad \text{where } k = (b-ac)/a.$$

If $k = r^2 > 0$ the solution is

$$\frac{1}{ar}\tan^{-1}\frac{y}{r} = \frac{1}{a}\sqrt{\{a/(b-ac)\}}\tan^{-1}\sqrt{\{(ax^2+ac)/(b-ac)\}}$$

and if $k = -s^2 < 0$ the solution is

$$\frac{1}{2sa}\log\frac{y-s}{y+s} = \frac{1}{2a}\sqrt{\{a/(ac-b)\}}\log\frac{[\sqrt{(x^2+c)}-\sqrt{\{(ac-b)/a\}}]^2}{x^2+b/a}.$$

Next we evaluate $\displaystyle\int \frac{dx}{(ax^2+b)\sqrt{(x^2+c)}}$.

Write $y = x/\sqrt{(x^2+c)}$ so that

$$y^2 = \frac{x^2}{x^2+c} = 1-\frac{c}{x^2+c}, \quad y\frac{dy}{dx} = \frac{cx}{(x^2+c)^2}, \quad \text{and} \quad x^2 = \frac{cy^2}{1-y^2};$$

then

$$\int \frac{dx}{(ax^2+b)\sqrt{(x^2+c)}} = \int \frac{cx\,dx}{(x^2+c)^2}\frac{(x^2+c)^{\frac{3}{2}}}{x^3}\frac{x^2}{c(ax^2+b)}$$

$$= \int y\,dy\frac{1}{y^3}\frac{y^2}{acy^2-by^2+b},$$

whence

$$\int \frac{dx}{(ax^2+b)\sqrt{(x^2+c)}} = \int \frac{dy}{b-(b-ac)y^2}$$

and the solution is completed as above.

8.861. The general case of the integral of

$$(Lx+M)/(ax^2+2bx+c)\sqrt{(Ax^2+2Bx+C)},$$

where $ax^2+2bx+c$ has no real factors, is easily reducible to an integral of the form 8.86.

Write $x = (rt+s)/(1+t)$ and so

$$ax^2+2bx+c = \{a(rt+s)^2+2b(rt+s)(1+t)+c(1+t)^2\}/(1+t)^2.$$

The coefficient of t in the numerator is zero if

$$ars+b(r+s)+c = 0. \tag{i}$$

Similarly,

$$Ax^2+2Bx+C = \{A(rt+s)^2+2B(rt+s)(1+t)+C(1+t)^2\}/(1+t)^2$$

and the coefficient of t in the numerator is zero if

$$Ars+B(r+s)+C = 0. \tag{ii}$$

Equations (i) and (ii) will both be satisfied if r, s are the roots of

$$(aB-bA)\theta^2-(cA-aC)\theta+(bC-cB) = 0, \tag{iii}$$

for this gives

$$ars+b(r+s) = \{a(bC-cB)+b(cA-aC)\}/(aB-bA) = -c,$$

and similarly $\qquad Ars+B(r+s) = -C.$

Moreover

$$x = \frac{rt+s}{1+t} = r+\frac{s-r}{t+1} \quad \text{and so} \quad \frac{dx}{dt} = \frac{r-s}{(1+t)^2}$$

and $\qquad Lx+M = \{(Lr+M)t+Ls+M\}/(1+t).$

Therefore if r, s are the roots of equation (iii) the integral is transformed into one of the form

$$\int \frac{lt+m}{(\alpha t^2+\beta)\sqrt{(\gamma t^2+\delta)}}\,dt = \frac{1}{\sqrt{\gamma}}\int \frac{lt+m}{(\alpha t^2+\beta)\sqrt{(t^2+\delta/\gamma)}}\,dt,$$

which we have evaluated in 8.86.

It remains to prove that equation (iii) has (different) real roots; this requires

$$(cA-aC)^2-4(aB-bA)(bC-cB) > 0,$$

subject to the condition $ac-b^2 > 0$ since $ax^2+2bx+c = 0$ has no real roots.

Let α, β, γ be any three numbers, then

$$(\alpha+\gamma-2\beta)^2 \geqslant 0,$$

whence $\qquad \alpha^2+\gamma^2+2\alpha\gamma \geqslant 4\beta(\alpha+\gamma)-4\beta^2$

and so
$$(\alpha-\gamma)^2 \geqslant 4(\alpha\beta+\beta\gamma-\alpha\gamma-\beta^2)$$
$$= 4(\beta-\alpha)(\gamma-\beta).$$
Take A/a, B/b, C/c for α, β, γ and we have
$$\left(\frac{A}{a}-\frac{C}{c}\right)^2 \geqslant 4\left(\frac{B}{b}-\frac{A}{a}\right)\left(\frac{C}{c}-\frac{B}{b}\right),$$
$$\therefore \qquad (cA-aC)^2 \geqslant 4\frac{ac}{b^2}(aB-bA)(bC-cB),$$
and so, since $ac/b^2 > 1$, the condition for real roots is satisfied. To ensure that the equation (iii) has *two different* roots we require also that $aB-bA$ is not zero; if $aB-bA = 0$ the transformation fails (since it reduces to $x =$ constant) but in this case both $ax^2+2bx+c$ and $Ax^2+2Bx+C$ are of the form $\dfrac{1}{a}(ax+b)^2+$constant and so the integral is reduced by the substitution $ax+b = t$.

8.87. If we denote the integral of
$$(Lx+M)/(ax^2+2bx+c)\sqrt{(Ax^2+2Bx+C)}$$
by $F(x, c)$ it follows that
$$\frac{d}{dx}\{F(x,c)\} = \frac{Lx+M}{(ax^2+2bx+c)\sqrt{(Ax^2+2Bx+C)}}$$
and therefore
$$\frac{d^p}{dc^p}\left(\frac{d}{dx}F(x,c)\right) = (-1)^p p! \frac{Lx+M}{(ax^2+2bx+c)^{p+1}\sqrt{(Ax^2+2Bx+C)}}$$
Hence if we assume the
$$\frac{d^p}{dc^p}\left(\frac{d}{dx}F(x,c)\right) = \frac{d}{dx}\left(\frac{d^p}{dc^p}F(x,c)\right),$$
we have
$$\int \frac{Lx+L}{(ax^2+2bx+c)^{p+1}\sqrt{(Ax^2+2Bx+C)}}\,dx = \frac{(-1)^p}{p!}\frac{d^p}{dc^p}F(x,c)$$

8.88. The integral of $R(x,y)$, where $y = \sqrt{(Ax^2+2Bx+C)}$ and $R(x, y)$ is a rational function, $P(x, y)/Q(x, y)$.

By separating the odd and even powers of y in the polynomials $P(x,y)$, $Q(x,y)$ the rational function $R(x,y)$ may be reduced to $(A+By)/(C+Dy)$, where A, B, C, D are polynomials in x alone. Multiplying the numerator and denominator by $C-Dy$, $(A+By)/(C+Dy)$ takes the form $E+Fy$, where E and F are

rational functions in x, and this again is of the form $E+G/y$ where E and G are rational functions. Expressing G in partial fractions the integral of $R(x, y)$ is reduced to a sum of integrals of the forms

$$\int E\,dx,\quad \int \frac{x^m}{y}\,dx,\quad \int \frac{dx}{(x-p)^m y},\quad \text{and}\quad \int \frac{Lx+M}{(ax^2+2bx+c)^{p+1}y}\,dx,$$

each of which has already been evaluated.

8.89. In conclusion we shall examine some special cases in which the integral of a quadratic irrational may be evaluated directly by a special substitution.

Consider first the integral of $1/\sqrt{\{(a-x)(x-b)\}}$, where $a > b$. We seek for a transformation which will make $a-x$ and $x-b$ both squares simultaneously; this requires $a-x = u^2$, $x-b = v^2$, and so $a-b = u^2+v^2$, which is satisfied by $u = r\cos\theta$, $v = r\sin\theta$, provided $r^2 = a-b$. Thus the desired transformation is

$$x = b+v^2 = b+(a-b)\sin^2\theta = a-u^2$$
$$= a-(a-b)\cos^2\theta = a\sin^2\theta+b\cos^2\theta$$

and, if $0 < \theta < \frac{1}{2}\pi$, the integral becomes

$$\int \frac{1}{(a-b)\sqrt{(\sin^2\theta\cos^2\theta)}}\frac{dx}{d\theta}\,d\theta = 2\int d\theta = 2\theta = 2\sin^{-1}\sqrt{\left(\frac{x-b}{a-b}\right)}.$$

To evaluate the integral of $1/\sqrt{\{(a-x)(b-x)\}}$, $a > b$, we require $a-x = u^2$, $b-x = v^2$, whence $a-b = u^2-v^2$, which is satisfied by $u = \sqrt{(a-b)}\operatorname{ch}\theta$, $v = \sqrt{(a-b)}\operatorname{sh}\theta$, and so

$$x = a-u^2 = a-(a-b)\operatorname{ch}^2\theta = b\operatorname{ch}^2\theta-a\operatorname{sh}^2\theta.$$

The transformed integral is

$$-2\int d\theta = -2\theta = 2\log e^{-\theta} = 2\log(\operatorname{ch}\theta-\operatorname{sh}\theta)$$
$$= 2\log\{\sqrt{(a-x)}-\sqrt{(b-x)}\}.$$

Similarly, we can evaluate the integrals of $\sqrt{\{(a-x)/(x-b)\}}$ and $\sqrt{\{(x-a)/(x-b)\}}$ by the transformations $x = b\cos^2\theta+a\sin^2\theta$ and $x = b\operatorname{ch}^2\theta-a\operatorname{sh}^2\theta$ respectively. We find

$$\int \sqrt{\left(\frac{a-x}{x-b}\right)}\,dx = (a-b)\int 2\cos^2\theta\,d\theta = (a-b)\int (1+\cos 2\theta)\,d\theta$$
$$= \sqrt{\{(a-x)(x-b)\}}+(a-b)\sin^{-1}\sqrt{\left(\frac{x-b}{a-b}\right)}$$

and

$$\int \sqrt{\left(\frac{x-a}{x-b}\right)}\,dx = \log\{\sqrt{(x-b)}-\sqrt{(x-a)}\}-\sqrt{\{(x-a)(x-b)\}}.$$

Example. To evaluate

$$\int \frac{x-1}{(2x^2-6x+5)\sqrt{(7x^2-22x+19)}}\,dx.$$

We make the substitution $x = \dfrac{rt+s}{t+1}$, where r and s are determined by the equations

$$2rs-3(r+s)+5 = 0,$$
$$7rs-11(r+s)+19 = 0,$$

whence $r = 1$, $s = 2$ (or vice versa).

From $x = \dfrac{t+2}{t+1} = 1+\dfrac{1}{t+1}$ we have $\dfrac{dx}{dt} = -\dfrac{1}{(t+1)^2}$ and

$$2x^2-6x+5 = \frac{1}{(t+1)^2}(t^2+1),$$

$$7x^2-22x+19 = \frac{1}{(t+1)^2}(4t^2+3),$$

and therefore

$$\int \frac{x-1}{(2x^2-6x+5)\sqrt{(7x^2-22x+19)}}\,dx = -\frac{1}{2}\int \frac{1}{(t^2+1)\sqrt{(t^2+\frac{3}{4})}}\,dt.$$

Take $y = \dfrac{t}{\sqrt{(t^2+\frac{3}{4})}}$ so that $y^2 = \dfrac{t^2}{t^2+\frac{3}{4}} = 1-\dfrac{\frac{3}{4}}{t^2+\frac{3}{4}}$ and therefore

$$y\frac{dy}{dt} = \frac{\frac{3}{4}t}{(t^2+\frac{3}{4})^2}.$$

Also $t^2+\frac{3}{4} = 3/4(1-y^2)$ and $t^2+1 = (4-y^2)/4(1-y^2)$, whence

$$\int \frac{1}{(1+t^2)\sqrt{(t^2+\frac{3}{4})}}\,dt = \int \frac{\frac{3}{4}t\,dt}{(t^2+\frac{3}{4})^2}\,\frac{\sqrt{(t^2+\frac{3}{4})}}{\frac{3}{4}t}\,\frac{t^2+\frac{3}{4}}{t^2+1}$$

$$= \int y\,dy\frac{1}{\frac{3}{4}y}\,\frac{3}{4-y^2} = \int \frac{4}{4-y^2}\,dy$$

$$= \int \frac{1}{2-y}\,dy+\int \frac{1}{2+y}\,dy = \log\frac{y+2}{y-2}$$

and therefore

$$\int \frac{x-1}{(2x^2-6x+5)\sqrt{(7x^2-22x+19)}}\,dx = \frac{1}{2}\log\frac{y-2}{y+2}.$$

IX

THE DEFINITE INTEGRAL

THE FIRST MEAN-VALUE THEOREM. CHANGE OF VARIABLE. GENERALIZED INTEGRALS. INTEGRAL OF A POWER SERIES

9. If two functions $f(x)$ and $g(x)$ differ only by a constant then for any x and y the difference $f(x)-f(y)$ is unchanged if we replace the function f by the function g, for if $f(x) = g(x)+c$ then

$$f(x)-f(y) = g(x)+c-(g(y)+c) = g(x)-g(y).$$

We have seen that an integral of a function is definite apart from an additive constant, accordingly if $F(x)$ is any integral of a function $f(x)$ then the value of $F(x)-F(y)$ is the same whatever integral $F(x)$ we choose; for instance x^2 is an integral of $2x$ and the difference x^2-y^2 is unchanged by replacing x^2 by any other integral of $2x$, such as x^2+3. Though independent of F in the sense we have described, the difference $F(x)-F(y)$ depends of course upon the values of x and y.

If $F(t)$ is uniformly differentiable in an interval (x, y), with derivative $f(t)$, then the difference $F(y)-F(x)$ is called the *definite integral* of $f(t)$ from x to y, and x, y are called the *limits*† of the integral.

We shall denote the difference $F(y)-F(x)$ for brevity by $[F(t)]_x^y$, where the t in $F(t)$ may be replaced by any other letter (including either x or y). The expression $\left[\int f(t)\,dt\right]_x^y$ will be further abbreviated to $\int\limits_x^y f(t)\,dt$.

Thus if $F(t)$ is any integral of $f(t)$ so that $F'(t) = f(t)$ in (x, y) then

$$\int\limits_x^y f(t)\,dt = F(y)-F(x).$$

In particular

$$\int\limits_x^x f(t)\,dt = 0, \quad \text{and} \quad \int\limits_y^x f(t)\,dt = -\int\limits_x^y f(t)\,dt.$$

† There is no relation between this use of the word 'limit' and the limit of a convergent sequence.

EXAMPLES. Since

$$\int \frac{1}{t}\,dt = \log t, \quad t > 0,$$

therefore $$\int_2^x \frac{1}{t}\,dt = \log x - \log 2 \quad (\text{provided } x > 0).$$

From $$\int \frac{1}{\sqrt{(1-t^2)}}\,dt = \sin^{-1}t, \quad -1 < t < 1,$$

we deduce

$$\int_{\frac{1}{3}}^{\frac{1}{2}} \frac{1}{\sqrt{(1-t^2)}}\,dt = \sin^{-1}\tfrac{1}{2} - \sin^{-1}\tfrac{1}{3} = \tfrac{1}{6}\pi - \cdot 3398 = \cdot 1838$$

and from

$$\int \frac{1}{\sqrt{(x^2-4)}}\,dx = \log\{x+\sqrt{(x^2-4)}\}, \quad x > 2,$$

we have

$$\int_3^4 \frac{1}{\sqrt{(x^2-4)}}\,dx = \log(4+\sqrt{12}) - \log(3+\sqrt{5}) = 2\cdot 0101 - 1\cdot 6556$$
$$= \cdot 3545.$$

9.01. The expression $[F(x)]_a^b$ satisfies the following simple theorems:

(1) If $F(x)$, $G(x)$ are equal apart from an additive constant then
$$[F(x)]_a^b = [G(x)]_a^b.$$
(2) $[F(x)]_a^b + [F(x)]_b^c = [F(x)]_a^c$ for any $F(x)$, a, b, and c.
(3) $[F(x)+G(x)]_a^b = [F(x)]_a^b + [G(x)]_a^b$ for any $F(x)$, $G(x)$, a, and b.

The proofs of these are obvious.

9.02. Each of the equations

$$\int_a^x f(t)\,dt = F(x) - F(a),$$
$$f(x) = F'(x)$$

is a consequence of the other, by definition.

9.021. $$\int_a^b -f(x)\,dx = -\int_a^b f(x)\,dx = \int_b^a f(x)\,dx,$$

for if $DF(x) = f(x)$ then $D(-F(x)) = -f(x)$ and so

$$\int_a^b -f(x)\,dx = [-F(x)]_a^b = F(a)-F(b) = \int_b^a f(x)\,dx$$

$$= -[F(b)-F(a)] = -\int_a^b f(x)\,dx.$$

9.03. From Theorem 8.2 and equation (3) of § 9.01 we have

$$\int_a^b \{f(x)+g(x)\}\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

9.04. From Theorem 8.4 and equations (1) and (3) of § 9.01 we have

$$\int_a^b f'(x)g(x)\,dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)\,dx.$$

9.1. If $f(x) > 0$ then $\int_a^x f(x)\,dx > 0$ for any $x > a$.

For $\dfrac{d}{dx}\left\{\int_a^x f(t)\,dt\right\} = f(x) > 0$ and so, by Theorem 3.62, $\int_a^x f(t)\,dt$ is steadily increasing; but $\int_a^a f(t)\,dt = 0$ and so $\int_a^x f(t)\,dt > 0$ when $x > a$.

Similarly, if $f(x) \geqslant 0$ then $\int_a^x f(t)\,dt \geqslant 0$, when $x \geqslant a$.

Furthermore, if $f(x) \geqslant 0$ in (a, b) and $f(x) > 0$ in a part of (a, b) then $\int_a^b f(t)\,dt > 0$; for if (c, d) is the part in which $f(x) > 0$ then $\int_c^d f(t)\,dt > 0$ and $\int_a^c f(t)\,dt \geqslant 0$, $\int_d^b f(t)\,dt \geqslant 0$ and so

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^d f(t)\,dt + \int_d^b f(t)\,dt > 0.$$

9.101. Since $|f(x)|-f(x) \geqslant 0$ it follow from 9.1 that

$$\int_a^b \{|f(x)|-f(x)\}\,dx \geqslant 0$$

and so, by 9.03,

$$\int_a^b |f(x)|\,dx - \int_a^b f(x)\,dx \geqslant 0,$$

i.e. $$\int_a^b |f(x)|\,dx \geqslant \int_a^b f(x)\,dx.$$

Similarly, since $|f(x)|+f(x) \geqslant 0$, therefore

$$\int_a^b |f(x)|\,dx \geqslant -\int_a^b f(x)\,dx.$$

Hence, since

$$\left|\int_a^b f(x)\,dx\right| = \int_a^b f(x)\,dx \quad \text{or} \quad -\int_a^b f(x)\,dx,$$

therefore $$\int_a^b |f(x)|\,dx \geqslant \left|\int_a^b f(x)\,dx\right|.$$

9.11. If $m < f(x) < M$ for any x in (a, b) then

$$m(b-a) < \int_a^b f(x)\,dx < M(b-a).$$

For $M-f(x) > 0$ in (a, b) and therefore $\int_a^b \{M-f(x)\}\,dx > 0$.

But

$$\int_a^b \{M-f(x)\}\,dx = \int_a^b M\,dx - \int_a^b f(x)\,dx$$

$$= [Mx]_a^b - \int_a^b f(x)\,dx = M(b-a) - \int_a^b f(x)\,dx$$

and therefore $M(b-a) > \int_a^b f(x)\,dx$. Similarly

$$\int_a^b f(x)\,dx > m(b-a).$$

9.12. If $m \leqslant f(x) \leqslant M$ for any x in (a, b) then

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a).$$

(Proof as in 9.11.)

9.121. From 9.11 and 9.12 it follows that if $m \leqslant f(x) \leqslant M$ in (a, b) and $m < f(x) < M$ in some part of (a, b) then

$$m(b-a) < \int_a^b f(x)\,dx < M(b-a).$$

9.13. If $f(x)$ is continuous and $m \leqslant f(x) \leqslant M$ in (a, b), m and M both being values of $f(x)$ in (a, b), then there is a number c between a and b such that

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(c).$$

For, by 9.12, $\int_a^b f(x)\,dx$ lies between $m(b-a)$ and $M(b-a)$ and so $\frac{1}{b-a}\int_a^b f(x)\,dx$ lies between m and M; but m and M are values of $f(x)$ in (a, b), and therefore $f(x)$, being continuous, attains the value $\frac{1}{b-a}\int_a^b f(x)\,dx$ at some point c in (a, b).

Theorem 9.13 may also be proved without assuming that we can find the least and greatest values of $f(x)$, viz. m, M. For if $F(x) = \int_a^x f(t)\,dt$, so that $F'(x) = f(x)$ in (a, b), then, by Theorem 3.61, we can find c_1 and c_2 in (a, b) such that

$$F'(c_1) \leqslant \frac{F(b)-F(a)}{b-a} \leqslant F'(c_2),$$

i.e.
$$f(c_1) \leqslant \left\{\int_a^b f(t)\,dt\right\}\Big/(b-a) \leqslant f(c_2);$$

but $f(x)$ is continuous and therefore attains the value

$$\left(\int_a^b f(t)\,dt\right)\Big/(b-a),$$

between $f(c_1)$ and $f(c_2)$, at a point c in (c_1, c_2).

Theorems 9.11, 9.12, 9.121, and 9.13 collectively are known as the *mean-value theorem* for integrals.

9.131. If $g(x)$ lies between $f(x)$ and $h(x)$ then $\int_\alpha^\beta g(x)\,dx$ lies between $\int_\alpha^\beta f(x)\,dx$ and $\int_\alpha^\beta h(x)\,dx$ when $f(x)-h(x)$ is of constant sign.

Suppose first that $\alpha < \beta$. If $f(x) < g(x) < h(x)$, then

$$g(x)-f(x) > 0 \quad \text{and} \quad h(x)-g(x) > 0$$

so that, by 9.1,

$$\int_{\alpha}^{\beta} \{g(x)-f(x)\}\,dx > 0, \qquad \int_{\alpha}^{\beta} \{h(x)-g(x)\}\,dx > 0,$$

whence
$$\int_{\alpha}^{\beta} f(x)\,dx < \int_{\alpha}^{\beta} g(x)\,dx < \int_{\alpha}^{\beta} h(x)\,dx.$$

Similarly, if $f(x) > g(x) > h(x)$ then

$$\int_{\alpha}^{\beta} f(x)\,dx > \int_{\alpha}^{\beta} g(x)\,dx > \int_{\alpha}^{\beta} h(x)\,dx.$$

If $\alpha > \beta$, then $f(x) < g(x) < h(x)$ implies

$$\int_{\beta}^{\alpha} f(x)\,dx < \int_{\beta}^{\alpha} g(x)\,dx < \int_{\beta}^{\alpha} h(x)\,dx,$$

so that, changing the sign throughout,

$$\int_{\alpha}^{\beta} f(x)\,dx > \int_{\alpha}^{\beta} g(x)\,dx > \int_{\alpha}^{\beta} h(x)\,dx$$

and similarly $f(x) > g(x) > h(x)$ implies

$$\int_{\alpha}^{\beta} f(x)\,dx < \int_{\alpha}^{\beta} g(x)\,dx < \int_{\alpha}^{\beta} h(x)\,dx,$$

which completes the proof.

9.14. If $f(x)$ is continuous in (a, b), and if in any part (c, d) of (a, b) we can determine points μ, ν such that $f(\mu) \leqslant f(x) \leqslant f(\nu)$ for any x in (c, d) then in each sub-interval (a_r, a_{r+1}) of (a, b) we can determine a point c_r such that

$$\int_{a}^{b} f(x)\,dx = f(c_0)(a_1-a_0)+f(c_1)(a_2-a_1)+\ldots+f(c_k)(a_{k+1}-a_k),$$

where $a_0 = a$ and $a_{k+1} = b$.

For by repeated application of 9.01 (2)

$$\int_{a}^{b} f(x)\,dx = \int_{a_0}^{a_{k+1}} f(x)\,dx = \int_{a_0}^{a_1} f(x)\,dx + \int_{a_1}^{a_2} f(x)\,dx + \ldots + \int_{a_k}^{a_{k+1}} f(x)\,dx.$$

By 9.13 there is a point c_r such that $\int_{a_r}^{a_{r+1}} f(x)\,dx = f(c_r)(a_{r+1}-a_r)$ and the required result follows.

9.141. $f(x)$ is continuous in (a, b) and the points (a_k) divide the interval (a, b) into a finite number of sub-intervals, each of length $0(n)$. If x_k is any point in the sub-interval (a_k, a_{k+1}) and if σ_n denotes the sum $\sum f(x_k)(a_{k+1}-a_k)$ taken over all the intervals of the subdivision, then $\sigma_n \to \int_a^b f(x)\,dx$.

Proof. Since $f(x)$ is continuous in (a, b) we can choose n so that, for a given p, $f(x)-f(x^*) = 0(p)$ for any two points x, x^* in (a, b), satisfying $x-x^* = 0(n)$. By 9.14, in each interval (a_k, a_{k+1}) there is a point c_k such that $\int_a^b f(x)\,dx = \sum f(c_k)(a_{k+1}-a_k)$, and as c_k and x_k both lie in (a_k, a_{k+1}), we have $c_k-x_k = 0(n)$, and therefore

$$\begin{aligned}\int_a^b f(x)\,dx - \sigma_n &= \sum f(c_k)(a_{k+1}-a_k) - \sum f(x_k)(a_{k+1}-a_k)\\ &= \sum \{f(c_k)-f(x_k)\}(a_{k+1}-a_k)\\ &= 0(p)\sum(a_{k+1}-a_k) = 0(p)(b-a),\end{aligned}$$

which proves that $\sigma_n \to \int_a^b f(x)\,dx$.

Theorem 9.141 enables a definite integral to be evaluated without appealing to the indefinite integral. For instance, if we divide $(0, 1)$ into m equal parts by the points $1/m,\ldots,(m-1)/m$, and form the sum $\sum_1^{m-1} \frac{r^3}{m^3}\left(\frac{r+1}{m}-\frac{r}{m}\right)$, the limit of this sum will equal $\int_0^1 x^3\,dx$.

But

$$\sum \frac{r^3}{m^3}\left(\frac{r+1}{m}-\frac{r}{m}\right) = \frac{1}{m^4}\sum r^3 = \frac{1}{m^4}\left[\frac{m(m-1)}{2}\right]^2 = \frac{1}{4}\left(1-\frac{1}{m}\right)^2,$$

which tends to $\frac{1}{4}$; thus $\int_0^1 x^3\,dx = \frac{1}{4}$, which may of course be verified by means of the indefinite integral $\int x^3\,dx = x^4/4$.

(To justify the application of 9.14 to the function x^3 we observe that in any part (c, d) of $(0, 1)$ we have $c^3 \leqslant x^3 \leqslant d^3$ for any x in (c, d).)

9.2. Change of variable in a definite integral

If (1) $\phi(t_0) = a$, $\phi(t_1) = b$;

(2) $\phi(t)$ is differentiable in (t_0, t_1);

(3) $f\{\phi(t)\}$ is continuous in (t_0, t_1);

then

9.21.
$$\int_a^b f(x)\,dx = \int_{t_0}^{t_1} f\{\phi(t)\}\phi'(t)\,dt.$$

By Theorem 2 54 it follows from conditions (2) and (3) that $f(x)$ is continuous in the interval $(a, \phi(t))$ for any t in (t_0, t_1); hence if

$$\int_a^x f(x)\,dx = F(x) \quad \text{then} \quad \int_a^{\phi(t)} f(x)\,dx = F\{\phi(t)\}$$

and therefore

$$\frac{d}{dt}\int_a^{\phi(t)} f(x)\,dx = F'\{\phi(t)\}\phi'(t) = f\{\phi(t)\}\phi'(t).$$

But
$$\frac{d}{dt}\int_{t_0}^{t} f\{\phi(t)\}\phi'(t)\,dt = f\{\phi(t)\}\phi'(t)$$

and so $\int_{t_0}^{t} f\{\phi(t)\}\phi'(t)\,dt$ and $\int_a^{\phi(t)} f(x)\,dx$, having the same derivative, differ only by a constant, and since each integral vanishes when $t = t_0$, this constant is zero and we have, for any t in (t_0, t_1)

$$\int_a^{\phi(t)} f(x)\,dx = \int_{t_0}^{t} f\{\phi(t)\}\phi'(t)\,dt.$$

Taking $t = t_1$ we have the stated result.

Condition (3) is necessary only if, for a t between t_0 and t_1, $\phi(t)$ takes a value outside (a, b), otherwise the continuity of $f\{\phi(t)\}$ follows from that of $f(x)$ in (a, b) and $\phi(t)$ in (t_0, t_1).

The theorem remains valid if we replace condition (3) by the condition that $f(x)$ is continuous in an interval (α, β) such that $\alpha \leqslant \phi(t) \leqslant \beta$ for any t in (t_0, t_1), since this condition, together with (2) above, implies the continuity of $f\{\phi(t)\}$ in (t_0, t_1).

EXAMPLES. Evaluate $\int_{(}^{1} \frac{2x}{x^2+1}\,dx$ by the substitution $x = \tan t$.

As t varies from 0 to $\frac{1}{4}\pi$, x varies from 0 to 1 and $\tan t$ is differentiable in $(0, \frac{1}{4}\pi)$, and so by equation 9.21

$$\int_0^1 \frac{2x}{x^2+1}\,dx = \int_0^{\frac{1}{4}\pi} \frac{2\tan t}{1+\tan^2 t}\sec^2 t\,dt = 2\int_0^{\frac{1}{4}\pi} \tan t\,dt$$

$$= [2\log\sec t]_0^{\frac{1}{4}\pi} = \log 2 - \log 1 = \log 2 = \cdot 69315.$$

Equation 9.21 may of course be used to derive either one of the integrals from the other. In the above example we have proceeded from the left-hand side of 9.21 to the right-hand side. To illustrate the converse process we shall evaluate the same integral $\int_0^1 \frac{2t}{t^2+1}\,dt$ by the transformation $x = t^2+1$. We have

$$\int_0^1 \frac{2t}{t^2+1}\,dt = \int_0^1 \frac{1}{t^2+1}\frac{d}{dt}(t^2+1)\,dt = \int_1^2 \frac{1}{x}\,dx = \log 2$$

as before. To justify this transformation we need but to observe that t^2+1 is differentiable in $(0, 1)$.

It is most important to ensure, in working examples, that any transformation used satisfies the conditions of Theorem 9.2 *throughout the range of integration.* For instance, if we seek to evaluate $\int_{\frac{1}{2}}^{\sqrt{3}/2} \frac{x}{\sqrt{(1-x^2)}}\,dx$ by the substitution $x = \sin t$ we might be tempted to write

$$\int_{\frac{1}{2}}^{\sqrt{3}/2} \frac{x}{\sqrt{(1-x^2)}}\,dx = \int_{\frac{5}{6}\pi}^{\frac{1}{3}\pi} \frac{\sin t}{\cos t}\cos t\,dt,$$

since $\sin\frac{1}{3}\pi = \sqrt{3}/2$, $\sin\frac{5}{6}\pi = \sin\frac{1}{6}\pi = \frac{1}{2}$,

$$= [-\cos t]_{\frac{5}{6}\pi}^{\frac{1}{3}\pi} = -\frac{1+\sqrt{3}}{2},$$

which is *false* because

$$\int_{\frac{1}{2}}^{\sqrt{3}/2} \frac{x}{\sqrt{(1-x^2)}}\,dx = [-\sqrt{(1-x^2)}]_{\frac{1}{2}}^{\sqrt{3}/2} = \frac{\sqrt{3}-1}{2}.$$

The explanation of the fallacy lies in the fact that $\frac{\sin t}{\sqrt{(1-\sin^2 t)}}$ is *not* continuous in the interval $(\frac{1}{3}\pi, \frac{5}{6}\pi)$ since the denominator

vanishes at the point $t = \frac{1}{2}\pi$ in this interval. Another false step lay in taking $\sqrt{(1-\sin^2 t)}$ as $\cos t$ throughout the interval $(\frac{1}{6}\pi, \frac{2}{3}\pi)$ whereas in the part of this interval from $\frac{1}{2}\pi$ to $\frac{2}{3}\pi$ $\cos t$ is negative and equal to $-\sqrt{(1-\sin^2 t)}$. It is, however, correct to write

$$\int_{\frac{1}{2}}^{\sqrt{3}/2} \frac{x}{\sqrt{(1-x^2)}}\,dx = \int_{\frac{1}{6}\pi}^{\frac{1}{3}\pi} \sin t\,dt.$$

The condition that $\phi(t)$ be differentiable throughout (t_0, t_1) in making the transformation $x = \phi(t)$ in the integral $\int_{t_0}^{t_1} f\{\phi(t)\}\phi'(t)\,dt$ shows that e.g. such an integral as $\int_{\frac{1}{2}\pi}^{\frac{3}{2}\pi} R(\sin t, \cos t)\,dt$ *cannot* be evaluated by the substitution $x = \tan\frac{1}{2}t$ since $\tan\frac{1}{2}t$ is *not* differentiable in $(\frac{1}{2}\pi, \frac{3}{2}\pi)$. In the use of transformations involving the inverse circular functions great care must be taken to introduce the inverse functions appropriate to the range of integration; thus if the range of integration is $0 \leqslant x \leqslant \pi$ we cannot use the transformation $x = \sin^{-1}t$ since the values of $\sin^{-1}t$ lie between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$, and if the range is $-\frac{1}{2}\pi < x < \frac{1}{2}\pi$ we may use $x = \sin^{-1}t$ or $x = \tan^{-1}t$ but not $x = \cos^{-1}t$.

A similar situation arises if we attempt to evaluate the integral $\int_{-1}^{2} x^2\,dx$ by the transformation $x^2 = y$, for in the interval

$$-1 \leqslant x < 0, \qquad x = -\sqrt{y},$$

and in the interval $0 < x \leqslant 2$, $x = +\sqrt{y}$, and so in these intervals $\frac{dx}{dy} = -\frac{1}{2\sqrt{y}}$ and $\frac{dx}{dy} = +\frac{1}{2\sqrt{y}}$ respectively, but the transformation function is *not* differentiable in the whole interval $(-1, 2)$ (since $\pm 1/2\sqrt{y}$ has no value for $y = 0$), and so the transformation $x^2 = y$ is not valid in this interval; correspondingly, if we seek to treat the integral $\int_{-1}^{2} x^2\,dx$ as the *result* of the transformation $y = x^2$ on the integral $\frac{1}{2}\int_{1}^{4} \sqrt{y}\,dy$ we find *either* that

$$\frac{1}{2}\int_{1}^{4} \sqrt{y}\,dy = \frac{1}{2}\int_{1}^{2} x\,2x\,dx = \int_{1}^{2} x^2\,dx$$

or $$\tfrac{1}{2}\int_1^4 \sqrt{y}\,dy = \tfrac{1}{2}\int_{-1}^{-2} -x\,2x\,dx = \int_1^2 t^2\,dt,$$

but neither $\tfrac{1}{2}\int_1^4 \sqrt{y}\,dy = \tfrac{1}{2}\int_{-1}^{2} 2x^2\,dx$ nor $\tfrac{1}{2}\int_1^4 \sqrt{y}\,dy = \tfrac{1}{2}\int_{-1}^{2} -2x^2\,dx$ is true, for $\sqrt{x^2} = +x$ if $x \geqslant 0$ and $\sqrt{x^2} = -x$ if $x \leqslant 0$, but neither $\sqrt{x^2} = +x$ nor $\sqrt{x^2} = -x$ is true throughout an interval which contains both negative and positive values of x.

This example illustrates also a further point. In applying the transformation $y = x^2$ to the integral $\int_1^4 \sqrt{y}\,dy$ it may be thought that there is some ambiguity as to the limits of the transformed integral, since both $(+2)^2 = 4$ and $(-2)^2 = 4$ and so too $(+1)^2 = 1$ and $(-1)^2 = 1$. It is true that we have a choice for the limits, but for each chosen pair we must examine the nature of the transformation in the chosen interval; for instance, if we choose the values $x = -1$ and $x = +2$ for the limits of the transformed integral then the integral becomes $\int_{-1}^{2} \sqrt{x^2}\,2x\,dx$, but this *cannot* be replaced either by $\int_{-1}^{2} x\,2x\,dx$ or by $\int_{-1}^{2} -x\,2x\,dx$ and must be split up into

$$\int_{-1}^{0} \sqrt{x^2}\,2x\,dx + \int_0^2 \sqrt{x^2}\,2x\,dx = \int_{-1}^{0} -x\,2x\,dx + \int_0^2 x\,2x\,dx.$$

Writing $t = -x$ in the first of these integrals it becomes $\int_1^0 2t^2\,dt$ and therefore

$$\begin{aligned}\int_{-1}^{2} \sqrt{x^2}\,2x\,dx &= \int_1^0 2t^2\,dt + \int_0^2 2x^2\,dx \\ &= \int_1^0 2t^2\,dt + \int_0^2 2t^2\,dt \\ &= \int_1^2 2t^2\,dt.\end{aligned}$$

It will readily be verified that, whichever choice of limits we make, a correct interpretation of the function $\sqrt{x^2}$ in the chosen interval

will always lead us to an integral equal to the integral $\int_1^2 2t^2\,dt$. That this *must* be the case we know already from Theorem 9.2 and the fact that a definite integral has a *unique* value.

9.3. If a function $f(x)$ has an integral, then $f(x)$ is the uniform derivative of this integral and therefore $f(x)$ is *continuous*. Thus a necessary condition for $f(x)$ to have an integral is that $f(x)$ be continuous. In order that $f(x)$ have a definite integral from a to b there must be a function $F(x)$ of which $f(x)$ is the derivative *throughout* the interval (a, b), and accordingly $f(x)$ must be continuous throughout (a, b). The step from an integral to a definite integral may not always be possible. For instance, although $\int \frac{1}{x}\,dx = \log x$, it is *not* true that $\int_{-2}^{-3} \frac{1}{x}\,dx = [\log x]_{-2}^{-3}$, for the function $\log x$ is defined, and has the derivative $1/x$, only for values of x greater than zero; to evaluate $\int_{-2}^{-3} \frac{1}{x}\,dx$ we make the transformation $x = -t$ and the integral becomes $\int_2^3 -\frac{1}{t}.-1.dt = \int_2^3 \frac{1}{t}\,dt$, and since $\log t$ is defined, and has the derivative $1/t$ in the interval $(2, 3)$, the value of this integral is $[\log t]_2^3 = \log \frac{3}{2}$.

Again, from $\int \frac{1}{\sqrt{(1-x^2)}}\,dx = \sin^{-1}x$ we can deduce

$$\int_0^{\frac{1}{2}} \frac{1}{\sqrt{(1-x^2)}}\,dx = [\sin^{-1}x]_0^{\frac{1}{2}} = \tfrac{1}{6}\pi$$

but *not* $\int_0^1 \frac{1}{\sqrt{(1-x^2)}}\,dx = [\sin^{-1}x]_0^1$ even though $\sin^{-1}x$ is defined throughout $(0, 1)$, for $\sin^{-1}x$ is differentiable, and has the derivative $1/\sqrt{(1-x^2)}$ only for $|x| < 1$. Of course, since $\frac{1}{\sqrt{(1-x^2)}}$ is defined and is continuous only for $|x| < 1$, $1/\sqrt{(1-x^2)}$ has no definite integral

over a range which is outside the interval $[-1, 1]$, and so such an expression as $\displaystyle\int_{2}^{3} \frac{1}{\sqrt{(1-x^2)}}\,dx$ is meaningless.

As a further illustration consider the integral $\displaystyle\int_{-3}^{-2} \frac{1}{\sqrt{(x^2-1)}}\,dx$. The function $1/\sqrt{(x^2-1)}$ is continuous if $x < -1$, but it is the derivative of $\mathrm{ch}^{-1}x$ only if $x > 1$, and so we cannot say that the value of the integral is $[\mathrm{ch}^{-1}x]_{-3}^{-2}$. As in the case of the integral $\displaystyle\int_{-2}^{-3} \frac{1}{x}\,dx$ we make the transformation $x = -t$ and the integral becomes $\displaystyle\int_{3}^{2} \frac{1}{\sqrt{(t^2-1)}}\,.\,-1\,.\,dt = \int_{2}^{3} \frac{1}{\sqrt{(t^2-1)}}\,dt$, which, since the range of integration is now positive, has the value

$$[\mathrm{ch}^{-1}t]_2^3 = \log\frac{3+2\sqrt{2}}{2+\sqrt{3}} = \cdot 4458.$$

Although continuity is a necessary condition for a function to have an integral there are continuous functions which cannot be integrated in terms of rational functions, circular functions, exponential or logarithmic functions. An instance of such an integral is

$$\int_0^x \frac{1}{\sqrt{(1-k^2\sin^2 x)}}\,dx, \quad k < 1,$$

which is known as an elliptic integral.

9.31. Standard results

(i) $$\int_0^\pi \cos mx \cos nx\,dx = \tfrac{1}{2}\int_0^\pi \cos(m+n)x\,dx + \tfrac{1}{2}\int_0^\pi \cos(m-n)x\,dx$$

$$= \left[\frac{\sin(m+n)x}{2(m+n)}\right]_0^\pi + \left[\frac{\sin(m-n)x}{2(m-n)}\right]_0^\pi,$$

provided $m^2 \neq n^2$,

$$= 0.$$

Thus, if $m^2 \neq n^2$,

$$\int_0^\pi \cos mx \cos nx\,dx = 0.$$

If $m^2 = n^2 \neq 0$.

$$\int_0^\pi \cos mx \cos nx \, dx = \int_0^\pi \cos^2 mx \, dx$$

$$= \tfrac{1}{2}\int_0^\pi \{1+\cos 2mx\}\, dx = [\tfrac{1}{2}x]_0^\pi + \left[\frac{\sin 2mx}{4m}\right]_0^\pi = \tfrac{1}{2}\pi,$$

$\therefore$
$$\int_0^\pi \cos^2 mx \, dx = \tfrac{1}{2}\pi.$$

(ii)
$$\int_0^1 \frac{dx}{a^2+b^2x^2} = \left[\frac{1}{ab}\tan^{-1}\frac{bx}{a}\right]_0^1 = \frac{1}{ab}\tan^{-1}\frac{b}{a};$$

$$\int_0^1 \frac{dx}{a^2-b^2x^2} = \left[\frac{1}{2ab}\log\frac{a+bx}{a-bx}\right]_0^1 = \frac{1}{2ab}\log\frac{a+b}{a-b}, \quad \text{provided } a^2 \not> b^2.$$

(iii) The integral of $\cos^n x$ from 0 to $\tfrac{1}{2}\pi$.

By reduction formula (4) of § 8.528

$$\int_0^{\frac{1}{2}\pi} \cos^n x \, dx = \frac{1}{n}[\sin x \cos^{n-1}x]_0^{\frac{1}{2}\pi} + \frac{n-1}{n}\int \cos^{n-2}x \, dx$$

$$= \frac{n-1}{n}\int_0^{\frac{1}{2}\pi} \cos^{n-2}x \, dx.$$

Hence if n is an even whole number

$$\int^{\frac{1}{2}\pi} \cos^n x \, dx = \frac{n-1.n-3 \ldots 3.1}{n.n-2\ldots 4.2}\int^{\frac{1}{2}\pi} d$$

and if n is od

$$\int_0^{\frac{1}{2}\pi} \cos^n x \, dx = \frac{n-1.n-3\ldots 4.2}{n.n-2\ldots 5.3}\int_0^{\frac{1}{2}\pi} \cos x \, dx.$$

Thus if we denote by $n_!$ (read semi-factorial n) the product

$$n.n-2.n-4\ldots 4.2$$

if n is even, and the product $n.n-2.n-4\ldots 3.1$ if n is odd, then we have

$$\int_0^{\frac{1}{2}\pi} \cos^n x \, dx = \frac{(n-1)_!}{n_!}\tfrac{1}{2}\pi, \quad \text{if } n \text{ is even}$$

and
$$\int_0^{\frac{1}{2}\pi} \cos^n x \, d \quad \frac{(n-1)_!}{n_!}, \quad \text{if } n \text{ is odd}$$

The transformation $y = \frac{1}{2}\pi - x$ shows that

$$\int\limits_0^{\frac{1}{2}\pi} \sin^n x\,dx = -\int\limits_{\frac{1}{2}\pi}^{0} \cos^n x\,dx = \int\limits_0^{\frac{1}{2}\pi} \cos^n x\,dx.$$

(iv) The integral of $\sin^m x \cos^n x$ from 0 to $\frac{1}{2}\pi$.

We have

$$\int\limits_0^{\frac{1}{2}\pi} \sin^m x \cos^n x\,dx = \left[\frac{\sin^{m+1}x\cos^{n-1}x}{m+n}\right]_0^{\frac{1}{2}\pi} + \frac{n-1}{m+n}\int\limits_0^{\frac{1}{2}\pi} \sin^m x \cos^{n-2} x\,dx.$$

Hence if n is even

$$\int\limits_0^{\frac{1}{2}\pi} \sin^m x \cos^n x\,dx = \frac{n-1\,.\,n-3\ldots 3\,.\,1}{m+n\,.\,m+n-2\ldots m+4\,.\,m+2}\int\limits_0^{\frac{1}{2}\pi} \sin^m x\,dx$$

Therefore if m and n are both even the value of the integral is

$$\frac{(n-1)_!(m-1)_!}{(m+n)_!}\tfrac{1}{2}\pi$$

and if n is even but m is odd, the value is

$$\frac{(n-1)_!(m-1)_!}{(m+n)_!}.$$

If n is odd

$$\int\limits^{\frac{1}{2}\pi} \sin^m x \cos^n x\,dx = \frac{n-1\,.\,n-3\ldots 4\,.\,2}{m+n\,.\,m+n-2\ldots m+5\,.\,m+3}\int\limits^{\frac{1}{2}\pi} \sin^m x \cos x\,dx$$

$$= \frac{(n-1)_!}{m+n\,.\,m+n-2\ldots m+3\,.\,m+1}$$

$$= \frac{(n-1)_!(m-1)_!}{(m+n)_!}.$$

Thus $$\int\limits_0^{\frac{1}{2}\pi} \sin^m x \cos^n x\,dx = \delta\{(m-1)_!(n-1)_!/(m+n)_!\},$$

where $\delta = \frac{1}{2}\pi$ if m and n are both even but $\delta = 1$ otherwise.

If m and n are both odd the integral may be expressed in terms of factorials. For if $m = 2p+1$, $n = 2q+1$ then

$$(m-1)_! = 2p\,.\,2p-2\ldots 2 = 2^p(p!),$$

$$(n-1)_! = 2q\,.\,2q-2\ldots 2 = 2^q(q!),$$

and

$$(m+n)_! = 2(p+q+1)\,.\,2(p+q)\,.\,2(p+q-1)\ldots 2 = 2^{p+q+1}(p+q+1)!,$$

and therefore

$$\int_0^{\frac{1}{2}\pi} \sin^{2p+1}x \cos^{2q+1}x \, dx = \frac{(p!)(q!)}{(p+q+1)!} \frac{2^p . 2^q}{2^{p+q+1}} = \frac{1}{2} \frac{(p!)(q!)}{(p+q+1)!}.$$

(v) *The integrals of $f(\sin x)$ and $xf(\sin x)$ from 0 to π.*

$$\int_0^{\pi} f(\sin x) \, dx = \int_0^{\frac{1}{2}\pi} f(\sin x) \, dx + \int_{\frac{1}{2}\pi}^{\pi} f(\sin x) \, dx.$$

In the second integral write $t = \pi - x$ and we have

$$\int_{\frac{1}{2}\pi}^{\pi} f(\sin x) \, dx = -\int_{\frac{1}{2}\pi}^{0} f(\sin t) \, dt = \int_0^{\frac{1}{2}\pi} f(\sin x) \, dx,$$

whence

$$\int_0^{\pi} f(\sin x) \, dx = 2\int_0^{\frac{1}{2}\pi} f(\sin x) \, dx = 2\int_0^{\frac{1}{2}\pi} f(\cos t) \, dt, \quad t = \tfrac{1}{2}\pi - x.$$

For instance, if n is even, then $\sin^m x \cos^n x$ may be expressed as a function of $\sin x$, and therefore

$$\int_0^{\pi} \sin^m x \cos^n x \, dx = \{(m-1)_{\text{¡}}(n-1)_{\text{¡}}/(m+n)_{\text{¡}}\}2\delta, \quad n \text{ being even.}$$

Also

$$\int_0^{\pi} \sin^2 x \, dx = 2\int_0^{\frac{1}{2}\pi} \sin^2 x \, dx = 2\int_0^{\frac{1}{2}\pi} \cos^2 x \, dx$$

$$= \int_0^{\frac{1}{2}\pi} (\sin^2 x + \cos^2 x) \, dx = \int_0^{\frac{1}{2}\pi} 1 \, dx = \tfrac{1}{2}\pi.$$

Further

$$\int_0^{\pi} xf(\sin x) \, dx = -\int_{\pi}^{0} (\pi - t) f(\sin t) \, dt, \quad \text{writing } t = \pi - x$$

$$= \pi \int_0^{\pi} f(\sin x) \, dx - \int_0^{\pi} xf(\sin x) \, dx$$

and so

$$\int_0^{\pi} xf(\sin x) \, dx = \tfrac{1}{2}\pi \int_0^{\pi} f(\sin x) \, dx = \pi \int_0^{\frac{1}{2}\pi} f(\sin x) \, dx.$$

For example

$$\int_0^{\frac{1}{2}\pi} \frac{\sin x}{2-\sin^2 x}\,dx = \int_0^{\frac{1}{2}\pi} \frac{\sin x}{1+\cos^2 x}\,dx = \left[-\tan^{-1}(\cos x)\right]_0^{\frac{1}{2}\pi} = \tfrac{1}{4}\pi$$

and therefore $$\int_0^{\pi} \frac{x\sin x}{1+\cos^2 x}\,dx = \tfrac{1}{4}\pi^2.$$

9.32. There are further instances when it is possible to evaluate (in finite form) the definite integral of a function without a knowledge of an integral of the function. This possibility is illustrated again in the following example.

To evaluate $\int_0^{\pi} \log\sin x\,dx$.

We have seen (in 9.31) that

$$\int_0^{\pi} \log\sin x\,dx = 2\int_0^{\frac{1}{2}\pi} \log\sin x\,dx.$$

Since $\sin x = 2\sin\frac{1}{2}x\cos\frac{1}{2}x$ therefore

$$\int_0^{\pi} \log\sin x\,dx = \int_0^{\pi} \log 2\,dx + \int_0^{\pi} \log\sin\tfrac{1}{2}x\,dx + \int_0^{\pi} \log\cos\tfrac{1}{2}x\,dx$$

$$= I_1+I_2+I_3 \quad \text{(say)}.$$

Writing $x = 2t$ in I_2 we have

$$I_2 = 2\int_0^{\frac{1}{2}\pi} \log\sin t\,dt = \int_0^{\pi} \log\sin x\,dx,$$

and writing $x = \pi-2t$ in I_3,

$$I_3 = 2\int_0^{\frac{1}{2}\pi} \log\sin t\,dt = \int_0^{\pi} \log\sin x\,dx,$$

and so $$\int_0^{\pi} \log\sin x\,dx = \pi\log 2+2\int_0^{\pi} \log\sin x\,dx,$$

and therefore $$\int_0^{\pi} \log\sin x\,dx = -\pi\log 2.$$

9.4. The generalized integral

If the integral $\int_a^b f(x)\,dx$ exists, then $\int_y^x f(t)\,dt$ is continuous in x and continuous in y in the interval (a, b) and therefore, for any λ in $[a, b]$,

$$\int^{b-1/n} f(x)\,dx \to \int^{b} f(x)\,dx$$

and

$$\int_{a+1/n}^{\lambda} f(x)\,dx \to \int^{a} f(x)\,dx,$$

whence:

9.401. $$\int_{a+1/n}^{\lambda} f(x)\,dx + \int^{b-1/n} f(x)\,dx$$

$$\to \int_a^{\lambda} f(x)\,dx + \int_{\lambda}^{b} f(x)\,dx = \int_a^b f(x)\,dx.$$

9.402. If $f(x)$ is continuous in the interval $(a+1/n, b-1/n)$ for any n, and if, for a certain λ in $[a, b]$, $\int_{\lambda}^{b-1/n} f(x)\,dx \to u$ and $\int_{a+1/n}^{\lambda} f(x)\,dx \to v$, then for any other point λ^* in $[a, b]$ both the sequences

$$\int^{b-1/n} f(x)\,dx \quad \text{and} \quad \int_{a+1/n}^{\lambda^*} f(x)\,dx$$

are convergent and

$$\lim \int^{b-1/n} f(x)\,dx + \lim \int_{a+1/n}^{\lambda^*} f(x)\,dx = u+v.$$

Suppose $\lambda^* < \lambda$, then

$$\int_{\lambda^*}^{b-1/n} = \int_{\lambda^*}^{\lambda} + \int_{\lambda}^{b-1/n} f(x)\,dx \to \int_{\lambda^*}^{\lambda} f(x)\,dx + u$$

and

$$\int_{a+1/n}^{\lambda^*} = \int_{a+1/n}^{\lambda} - \int_{\lambda^*}^{\lambda} f(x)\,dx \to v - \int_{\lambda^*}^{\lambda} f(x)\,dx,$$

whence $$\lim \int_{a+1/n}^{\lambda^*} f(x)\,dx + \lim \int_{\lambda^*}^{b-1/n} f(x)\,dx = u+v$$

The sum of the limits

$$\lim \int_{a+1/n}^{\lambda} f(x)\,dx + \lim \int_{\lambda}^{b-1/n} f(x)\,dx$$

is called the *generalized integral* of $f(x)$ from a to b. The generalized integral is also denoted by $\int_a^b f(x)\,dx$, so that

$$\int_a^b f(x)\,dx = \lim \int_{a+1/n}^{\lambda} f(x)\,dx + \lim \int_{\lambda}^{b-1/n} f(x)\,dx$$

for any λ in $[a, b]$. This dual usage of the integral sign leads to no confusion because, by equation 9.401, when the integral exists in the ordinary sense, its value is the same as that of the generalized integral. If the generalized integrals $\int_a^b f(x)\,dx$ and $\int_b^c f(x)\,dx$, $a < b < c$ both exist then we *define*:

$$\int_{a}^{c} f(x)\,dx = \int_{a}^{b} f(x)\,dx + \int_{b}^{c} f(x)\,dx$$

9.41. If $f(x)$ is continuous in $(a, b-1/n)$ for any n, and i

$$\int_a^{b-1/n} f(x)\,dx \to g$$

then g is the value of the generalized integral from a to b. For if λ lies in $[a, b]$, then $f(x)$ is continuous in (a, λ) and so

$$\int_{a+1/n}^{\lambda} f(x)\,dx \to \int_a^{\lambda} f(x)\,dx.$$

Furthermore

$$\int_a^{b-1/n} f(x)\,dx = \int_a^{\lambda} f(x)\,dx + \int_{\lambda}^{b-1/n} f(x)\,dx,$$

and therefore

$$\lim \int_{a+1/n}^{\lambda} f(x)\,dx + \lim \int_{\lambda}^{b-1/n} f(x)\,dx = \int_a^{\lambda} f(x)\,dx + g - \int_a^{\lambda} f(x)\,dx = g$$

Similarly, if $f(x)$ is continuous in $(a+1/n,\, b)$ for any n, and if

$$\int_{a+1/n}^{b} f(x)\, dx \to g,$$

then g is the value of the generalized integral $\int_a^b f(x)\, dx$.

9.411. If the generalized integral $\int_a^b f(x)\, dx$ exists then

$$\int_{b-1/n}^{b} f(x)\, dx \to 0.$$

By 9.402, the generalized integral $\int_{b-1/n}^{b} f(x)\, dx$ exists for any fixed n. Since $\int_a^b f(x)\, dx$ exists, $\int_{\lambda}^{b-1/n} f(x)\, dx$ is convergent, and therefore

$$\left|\int_{b-1/n}^{b-1/N} f(x)\, dx\right| < \frac{1}{k} \quad \text{for } N > n \geqslant n_k,$$

whence

$$\left|\int_{b-1/n}^{b} f(x)\, dx\right| \leqslant \frac{1}{k}, \quad n \geqslant n_k,$$

i.e.

$$\int_{b-1/n}^{b} f(x)\, dx \to 0.$$

Similarly,

$$\int^{a+1/n} f(x)\, dx \to 0.$$

9.42. If both the generalized integrals $\int_a^b f(x)\, dx$ and $\int_a^b g(x)\, dx$ exist and if $f(x) \leqslant g(x)$, then

$$\int_a^b f(x)\, dx \leqslant \int_a^b g(x)\, dx.$$

It suffices to consider the case when $f(x)$ and $g(x)$ are both continuous in $(a,\, b-1/n)$, and $f(x) \leqslant g(x)$ in this interval, since the general case may be reduced to a sum of such cases. Then

$$\int_a^{b-1/n} f(x)\, dx \leqslant \int_a^{b-1/n} g(x)\, dx,$$

whence the result follows.

9.421. If $f(x)$ is continuous in $(a, b-1/n)$ for any n, and positive near b, and if $\int_a^b f(x)\,dx$ exists, then for any sequence b_n such that $b_n < b$ and $b_n \to b$ we have

$$\int_a^{b_n} f(x)\,dx \to \int_a^b f(x)\,dx.$$

Since $b_n \to b$, we can find n_k so that $b_n > b-1/k$ for $n \geqslant n_k$, and therefore, as $f(x)$ is positive near b, we have, for any p,

$$\int_{b_n}^{b-1/p} f(x)\,dx \leqslant \int_{b-1/k}^{b-1/p} f(x)\,dx$$

and therefore

$$\int_{b_n}^{b} f(x)\,dx \leqslant \int_{b-1/k}^{b} f(x)\,dx \to 0 \quad \text{by 9.411.}$$

Hence, using 9.402,

$$\int_a^{b_n} f(x)\,dx = \int_a^b f(x)\,dx - \int_{b_n}^b f(x)\,dx \to \int_a^b f(x)\,dx.$$

Similarly, if $f(x)$ is continuous in $(a+1/n, b)$, positive near a, and if $\int_a^b f(x)\,dx$ exists, then

$$\int_{a_n}^{b} f(x)\,dx \to \int_a^b f(x)\,dx, \quad \text{provided } a_n > a \text{ and } a_n \to a.$$

9.422. Conversely, if $f(x)$ is continuous in $(a, b-1/n)$ for any n, positive near b, and if $\int_a^{b_n} f(x)\,dx$ is convergent for a sequence b_n such that $b_n < b$ and $b_n \to b$ then $\int_a^b f(x)\,dx$ exists and is the limit of $\int_a^{b_n} f(x)\,dx$.

For if $1/n < b-b_k$ and $b-b_\lambda < 1/N$, then

$$\int_{b-1/n}^{b-1/N} f(x)\,dx \leqslant \int_{b_k}^{b_\lambda} f(x)\,dx < \frac{1}{p}, \quad \lambda > k \geqslant k_p,$$

since $\int_a^{b_n} f(x)\,dx$ converges, and therefore $\int_a^{b-1/n} f(x)\,dx$ converges.

Thus $\int_a^b f(x)\,dx$ exists. Hence by 9.421

$$\int_a^{b_n} f(x)\,dx \to \int_a^b f(x)\,dx.$$

The condition that $f(x)$ be positive near b may be replaced by the condition that it be negative, for if

$$\int_a^{b-1/n} (-f(x))\,dx = -\int_a^{b-1/n} f(x)\,dx$$

is convergent, then $\int_a^{b-1/n} f(x)\,dx$ is convergent.

9.423. We may omit the condition, in Theorem 9.421, that $f(x)$ is positive near b, if in addition to the other conditions upon $f(x)$ given in 9.421, the generalized integral $\int_a^b |f(x)|\,dx$ exists.

For, applying 9.421 to the positive function $|f(x)|$, it follows that

$$\int_{b_n}^b |f(x)|\,dx \to 0.$$

Hence since $|f(x)| \geqslant \pm f(x)$ in $(a, b]$, therefore by 9.42

$$\left|\int^b f(x)\,dx\right| \leqslant \int^b |f(x)|\,dx \to 0$$

so that

$$\int_a^{b_n} f(x)\,dx = \int_a^b f(x)\,dx - \int_{b_n}^b f(x)\,dx \to \int_a^b f(x)\,dx.$$

9.424. It is important to observe that the fundamental theorem for definite integrals

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

does not extend to generalized integrals, for when $f(x) = \dfrac{d}{dx}\displaystyle\int_a^x f(t)\,dt$ then $f(x)$, being a uniform derivative, is continuous in (a, x), which is not the case when $\int^x f(t)\,dt$ is a generalized integral.

9.425. If $f(x)$ is continuous in $(a, b-1/n)$ for any n, and if $(b-x)^\lambda f(x)$ is bounded in $(a, b]$, where $\lambda < 1$, then $\int_a^b f(x)\,dx$ exists.

For if $|(b-x)^\lambda f(x)| < M$, then for $N > n$,

$$\left|\int_{b-1/n}^{b-1/N} f(x)\,dx\right| \leqslant \int_{b-1/n}^{b-1/N} |f(x)|\,dx$$

$$\leqslant M \int_{b-1/n}^{b-1/N} (b-x)^{-\lambda}\,dx = M\left[-\frac{(b-x)^{1-\lambda}}{1-\lambda}\right]_{b-1/n}^{b-1/N}$$

$$= \frac{M}{1-\lambda}\left[\frac{1}{n^{1-\lambda}}-\frac{1}{N^{1-\lambda}}\right] < \frac{M}{1-\lambda}\frac{1}{n^{1-\lambda}} \to 0,$$

since $1-\lambda > 0$.

Thus $\int_a^{b-1/n} f(x)\,dx$ is convergent.

In particular $\int_a^b f(x)\,dx$ exists if $f(x)$ is bounded in $(a, b]$, for

$$\text{if } \lambda = 0, \qquad |(b-x)^\lambda f(x)| = |f(x)|.$$

Similarly, if $f(x)$ is continuous in $(a+1/n, b)$ for any n, and if $(x-a)^\lambda f(x)$ is bounded, then $\int_a^b f(x)\,dx$ exists.

9.43. If $f(x)$ is continuous in (a, n) for any n, and if $\int_a^n f(x)\,dx \to v$, then v is called *the generalized integral of* $f(x)$ *from* a *to infinity*. This generalized integral is denoted by

$$\int_a^\infty f(x)\,dx,$$

so that, by definition,

$$\int_a^\infty f(x)\,dx = \lim \int_a^n f(x)\,dx.$$

It is important to realize that the expression 'the integral from a to infinity' means neither more nor less than 'the limit of the integral from a to n', and in no way implies that there is some hitherto unmentioned number infinity. An analogous notation is often employed for series, the limit of the sum $\sum_1^n a_r$ being denoted by $\sum_1^\infty a_r$.

If $\int_{-n}^{b} f(x)\,dx$ is convergent and tends to w then w is called the *generalized integral from minus infinity to b*, and we write

$$\int_{-\infty}^{b} f(x)\,dx = \lim \int_{-n}^{b} f(x)\,dx.$$

If both the sequences $\int_{-n}^{a} f(x)\,dx$ and $\int_{a}^{n} f(x)\,dx$ are convergent we define

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim \int_{a}^{n} f(x)\,dx + \lim \int_{-n}^{a} f(x)\,dx.$$

9.44. If $f(x)$ is continuous, and if, for $p > 0$, $|x^{1+p}f(x)| < M$, in the interval (a, n), for any n, then the generalized integral $\int_{a}^{\infty} f(x)\,dx$ exists.

For if $N > n$, then

$$\left|\int_{a}^{N} f(x)\,dx - \int_{a}^{n} f(x)\,dx\right| = \left|\int_{n}^{N} f(x)\,dx\right| < M \int_{n}^{N} \frac{1}{x^{1+p}}\,dx$$

$$= \frac{M}{p}\left[-\frac{1}{x^{p}}\right]_{n}^{N} = \frac{M}{p}\left[\frac{1}{n^{p}} - \frac{1}{N^{p}}\right] < \frac{M}{p}\frac{1}{n^{p}} \to 0, \quad \text{since } p > 0.$$

Thus $\int_{a}^{n} f(x)\,dx$ is convergent, and so $\int_{a}^{\infty} f(x)\,dx$ exists.

Similarly, if $|x^{1+p}f(x)| < M$ in $(-n, a)$ for any n, the integral $\int_{-\infty}^{a} f(x)\,dx$ exists.

For if $N > n$,

$$\left|\int_{-N}^{a} f(x)\,dx - \int_{-n}^{a} f(x)\,dx\right| = \left|\int_{-N}^{-n} f(x)\,dx\right| < M \int_{-N}^{-n} \frac{1}{|x|^{1+p}}\,dx$$

$$= M \int_{n}^{N} \frac{1}{x^{1+p}}\,dx \to 0.$$

Accordingly, if $f(x)$ is continuous and $|x^{1+p}f(x)| < M$ in $(-n, n)$ for any n then $\int_{-\infty}^{\infty} f(x)\,dx$ exists.

9.45. If $f(x)$ is regular in (a, n) for any n, and $|f(x)| < M$ in (a, n), then $\int_a^\infty \frac{f'(x)}{x^p}\,dx$ exists when $p > 0$, $a > 0$.

For
$$\int_a^n \frac{f'(x)}{x^p}\,dx = \left[\frac{f(x)}{x^p}\right]_a^n + p\int_a^n \frac{f(x)}{x^{1+p}}\,dx.$$

Since $|f(x)| < M$, therefore $\left|x^{1+p}\frac{f(x)}{x^{1+p}}\right| < M$ so that, by 9.44, the second integral is convergent, and furthermore $\left|\frac{f(n)}{n^p}\right| < \frac{M}{n^p} \to 0$, and so
$$\int_a^n \frac{f'(x)}{x^p}\,dx \to -\frac{f(a)}{a^p} + p\int_a^\infty \frac{f(x)}{x^{1+p}}\,dx.$$

EXAMPLES. Since $\sin^{-1}x$ is differentiable, with derivative $1/\sqrt{(1-x^2)}$ in the interval $(0, 1-1/n)$, for any n (but *not* in the interval $(0, 1)$) we have
$$\int_0^{1-1/n} \frac{1}{\sqrt{(1-x^2)}}\,dx = \sin^{-1}\left(1-\frac{1}{n}\right) - \sin^{-1}0$$
$$\to \sin^{-1}1 - \sin^{-1}0,$$
since $\sin^{-1}x$ is continuous in $(-1, 1)$,
$$= \tfrac{1}{2}\pi$$

and therefore
$$\int_0^1 \frac{1}{\sqrt{(1-x^2)}}\,dx = \tfrac{1}{2}\pi.$$

For any x,
$$\int e^{-ax}\cos bx\,dx = e^{-ax}\{-a\cos bx + b\sin bx\}/(a^2+b^2),$$
and therefore
$$\int_0^n e^{-ax}\cos bx\,dx = e^{-an}\{-a\cos bn + b\sin bn\}/(a^2+b^2) + a/(a^2+b^2).$$

If a is positive,
$$|e^{-an}\cos bn| \leqslant e^{-an} < 1/an \to 0,$$
and similarly
$$e^{-an}\sin bn \to 0.$$

Therefore
$$\int_0^\infty e^{-ax}\cos bx\,dx = a/(a^2+b^2), \quad a > 0.$$

From $$\int_0^n x^p e^{-x}\,dx = [-x^p e^{-x}]_0^n + p\int_0^n x^{p-1}e^{-x}\,dx,$$
since $n^p/e^n \to 0$, we have (by 9.44)
$$\int_0^\infty x^p e^{-x}\,dx = p\int_0^\infty x^{p-1}e^{-x}\,dx$$
and therefore, if p is a positive integer,
$$\int_0^\infty x^p e^{-x}\,dx = p!\int_0^\infty e^{-x}\,dx.$$
But $$\int_0^n e^{-x}\,dx = [-e^{-x}]_0^n = 1-e^{-n} \to 1,$$
and so $$\int_0^\infty x^p e^{-x}\,dx = p!.$$

If $a^2 > b^2$ and $0 < x < \pi$, then
$$\int \frac{1}{a+b\cos x}\,dx = \frac{1}{\sqrt{(a^2-b^2)}}\cos^{-1}\left\{\frac{a\cos x+b}{a+b\cos x}\right\};$$
the equality does not extend to the points $x = 0$, $x = \pi$, since at these points $\dfrac{a\cos x+b}{a+b\cos x}$ takes the values 1 and -1 respectively, and $\cos^{-1}y$ is *differentiable* only in the open interval $[-1, 1]$. Hence
$$\int_{1/n}^{\frac{1}{2}\pi} \frac{1}{a+b\cos x}\,dx = \frac{1}{\sqrt{(a^2-b^2)}}\left[\cos^{-1}\left\{\frac{a\cos x+b}{a+b\cos x}\right\}\right]_{1/n}^{\frac{1}{2}\pi}$$
$$= \frac{1}{\sqrt{(a^2-b^2)}}\left\{\cos^{-1}\frac{b}{a} - \cos^{-1}\left(\frac{a\cos 1/n+b}{a+b\cos 1/n}\right)\right\}$$
$$\to \frac{1}{\sqrt{(a^2-b^2)}}\left(\cos^{-1}\frac{b}{a} - \cos^{-1}1\right),$$
since $\cos^{-1}y$ is *continuous* in the closed interval $(-1, 1)$.

Similarly,
$$\int_{\frac{1}{2}\pi}^{\pi-1/n} \frac{1}{a+b\cos x}\,dx \to \frac{1}{\sqrt{(a^2-b^2)}}\left(\cos^{-1}(-1) - \cos^{-1}\frac{b}{a}\right),$$

$$\int_0^\pi \frac{1}{a+b\cos x}\,dx = \frac{\pi}{\sqrt{(a^2-b^2)}}.$$

If $a^2 < b^2$, then $a+b\cos x$ vanishes for $\cos x = -\frac{a}{b}$, i.e.

$$\tan\tfrac{1}{2}x = \sqrt{\left(\frac{b+a}{b-a}\right)} = \mu \quad \text{(say).}$$

Consider the integral from 0 to $2\tan^{-1}(\mu-1/n)$; in this interval $\tan\frac{1}{2}x < \mu$ and so, writing $t = \tan\frac{1}{2}x$

$$\int^{2\tan^{-1}(\mu-1/n)} \frac{1}{a+b\cos x}\,dx = \frac{1}{\mu(a-b)}\left[\log\frac{\mu-t}{\mu+t}\right]_0^{\mu-1/n}$$

$$= -\frac{1}{\mu(a-b)}\left\{\log n + \log\left(2\mu-\frac{1}{n}\right)\right\},$$

which is not convergent since $\log n^2-\log n = \log n > 1$, for $n > 3$ Hence, since $a+b\cos x$ is of constant sign in $(0, 2\tan^{-1}\mu]$, it follow that the integral over this interval does not exist.

9.5. The integral of a power series

If $\sum a_n x^n$ is convergent in the interval $(0, X)$ and if $f(x)$ is it limit and $F(x)$ is an integral of $f(x)$, then $F'(x) = f(x)$ and so, b Theorem 3.71,

$$F(x)-F(0) = \sum a_n \frac{x^{n+1}}{n+1},$$

$$\int_0^x f(x)\,dx = a_0 x + a_1\frac{x^2}{2} + a_2\frac{x^3}{3} + \ldots \quad \text{for } x/X \leqslant$$

Since $a_n\dfrac{x^{n+1}}{n+1} = \displaystyle\int_0^x a_n x^n\,dx$ for any n this result may be expressed by saying that a power series may be integrated term by term in any interval in which it is convergent.

9.51. *Series for* $\tan^{-1}x$. Since

$$1-x^2+x^4-x^6+\ldots = 1/(1+x^2) \quad \text{provided } |x| < 1,$$

it follows from 9.5 that

$$\tan^{-1}x = \int_0^x \frac{1}{1+x^2}\,dx = x-\frac{x^3}{3}+\frac{x^5}{5}-\ldots \quad \text{for } |x| < 1.$$

Since $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\dots$ is convergent,† if $\phi(x)$ denotes the limit of $x-\dfrac{x^3}{3}+\dfrac{x^5}{5}-\dots$, then $\phi(x)$ is continuous in $(0, 1)$ and therefore

$$\phi\left(1-\frac{1}{n}\right) \to \phi(1);$$

but $\tan^{-1}x$ is also continuous (for any x) so that

$$\tan^{-1}\left(1-\frac{1}{n}\right) \to \tan^{-1}1 = \tfrac{1}{4}\pi.$$

Thus:

9.52. $$\tfrac{1}{4}\pi = 1-\tfrac{1}{3}+\tfrac{1}{5}-\tfrac{1}{7}+\dots.$$

Formula 9.52 is far more convenient for the calculation of the value of π than the method described in § 5.2, but even so, as many as 500 terms of the series are needed to obtain π to 3 places of decimals. We can however readily obtain, from 9.51, other series from which the expansion of π can be obtained with reasonable rapidity.

For instance, since

$$\tan\{\tan^{-1}\tfrac{1}{2}+\tan^{-1}\tfrac{1}{3}\} = (\tfrac{1}{2}+\tfrac{1}{3})/(1-\tfrac{1}{6}) = 1,$$

therefore

$$\begin{aligned}\tfrac{1}{4}\pi &= \tan^{-1}1 = \tan^{-1}\tfrac{1}{2}+\tan^{-1}\tfrac{1}{3}\\ &= \left(\frac{1}{2}+\frac{1}{3}\right)-\frac{1}{3}\left(\frac{1}{2^3}+\frac{1}{3^3}\right)+\frac{1}{5}\left(\frac{1}{2^5}+\frac{1}{3^5}\right)-\dots,\end{aligned}$$

and 50 terms of this series give π to 30 places of decimals.

Similarly,

$$\tan\left(2\tan^{-1}\frac{1}{5}\right) = \frac{5}{12} \quad\text{and so}\quad \tan\left(4\tan^{-1}\frac{1}{5}\right) = \frac{120}{119},$$

† If $s_n = 1-\dfrac{1}{3}+\dfrac{1}{5}-\dots+\dfrac{(-1)^{n+1}}{2n-1}$, then

$$\begin{aligned}|s_{n+p}-s_n| &= \frac{1}{2n+1}-\frac{1}{2n+3}+\dots+\frac{(-1)^{p-1}}{2n+2p-1}\\ &= \frac{1}{2n+1}-\left(\frac{1}{2n+3}-\frac{1}{2n+5}\right)-\dots < \frac{1}{2n+1},\end{aligned}$$

which proves that s_n is convergent.

This is a particular case of the theorem proved in Example 1.3.

whence

$$\tan\left\{4\tan^{-1}\frac{1}{5}-\tan^{-1}\frac{1}{239}\right\} = \left(\frac{120}{119}-\frac{1}{239}\right)\Big/\left(1+\frac{120}{119.239}\right)$$

$$= \frac{120.239-119}{119.239+120} = \frac{119.239+120}{119.239+120} = 1$$

and therefore

$$\tfrac{1}{4}\pi = \tan^{-1}1 = 4\tan^{-1}\frac{1}{5}-\tan^{-1}\frac{1}{239}$$

$$= 4\left\{\frac{1}{5}-\frac{1}{3}\cdot\frac{1}{5^3}+\frac{1}{5}\cdot\frac{1}{5^5}-\ldots\right\}-\left\{\frac{1}{239}-\frac{1}{3}\cdot\frac{1}{239^3}+\frac{1}{5}\cdot\frac{1}{239^5}-\ldots\right\};$$

50 terms of the first series and 15 of the second suffice to give π to 70 places.

9.53. *Series for* $\log(1+x)$. Integrating the series

$$\frac{1}{1+x} = 1-x+x^2-x^3+\ldots, \quad |x| < 1,$$

we obtain

$$\log(1+x) = x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots, \quad |x| < 1.$$

Since the series $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$ is convergent and $\log(1+x)$ is continuous in (0, 1) it follows as in 9.51 that

$$\log 2 = 1-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\ldots.$$

Similarly, integrating

$$\frac{1}{1-x} = 1+x+x^2+x^3+\ldots, \quad |x| < 1,$$

we have

$$-\log(1-x) = x+\frac{x^2}{2}+\frac{x^3}{3}+\frac{x^4}{4}+\ldots, \quad |x| < 1,$$

and so

$$\log\frac{1+x}{1-x} = \log(1+x)-\log(1-x) = 2\left\{x+\frac{x^3}{3}+\frac{x^5}{5}+\ldots\right\}, \quad |x| < 1.$$

By means of the identity

$$y = \left(1+\frac{y-1}{y+1}\right)\Big/\left(1-\frac{y-1}{y+1}\right)$$

we obtain the convergent series

$$\log y = 2\left\{\frac{y-1}{y+1}+\frac{1}{3}\left(\frac{y-1}{y+1}\right)^3+\frac{1}{5}\left(\frac{y-1}{y+1}\right)^5+\ldots\right\}, \quad \text{valid for } y > 0.$$

9.54. We conclude this chapter with an account of a simple method by which the expansions of $\tan^{-1}x$ and $\log(1+x)$ may be obtained, without integrating a power series.

Since

$$1+t+t^2+\ldots+t^{n-1} = (1-t^n)/(1-t), \quad \text{for all } t \text{ except } t = 1,$$

therefore

$$\frac{1}{1+t^2} = 1-t^2+t^4-\ldots+(-1)^{n-1}t^{2n-2}+(-1)^n\frac{t^{2n}}{1+t^2}$$

and so

$$\int_0^x \frac{1}{1+t^2}\,dt = \int_0^x \{1-t^2+t^4-\ldots+(-1)^{n-1}t^{2n-2}\}\,dt+(-1)^n\int_0^x \frac{t^{2n}}{1+t^2}\,dt,$$

that is,

$$\tan^{-1}x = x-\frac{x^3}{3}+\frac{x^5}{5}-\ldots+(-1)^{n-1}\frac{x^{2n-1}}{2n-1}+(-1)^n\int_0^x \frac{t^{2n}}{1+t^2}\,dt.$$

But $\dfrac{t^{2n}}{1+t^2} \leqslant t^{2n}$, and so, if $0 \leqslant x \leqslant 1$,

$$\int_0^x \frac{t^{2n}}{1+t^2}\,dt \leqslant \int_0^x t^{2n}\,dt \quad \frac{x^{2n+1}}{2n+1} \leqslant \frac{1}{2n+1};$$

and if $-1 \leqslant x \leqslant 0$, then

$$\left|\int_0^x \frac{t^{2n}}{1+t^2}\,dt\right| = \left|-\int_0^{-x} \frac{u}{1+}\,du\right|, \quad \text{where } u = -t,$$

$$= \int^{-x} \frac{u^{2n}}{1+u^2}\,du \leqslant \frac{(-x)^{2n+1}}{2n+1} \leqslant \frac{1}{2n+1}.$$

Accordingly

$$|x-\frac{x^3}{3}+\frac{x^5}{5}-\ldots+(-1)^{n-1}\frac{x^{2n-1}}{2n+1}-\tan^{-1}x \leqslant \frac{1}{2n+1},$$

provided $-1 \leqslant x \leqslant 1$,

which proves that the series

$$x-\frac{x^3}{3}+\frac{x^5}{5}\cdot$$

converges to $\tan^{-1}x$, provided $-1 \leqslant x \leqslant 1$, since $1/(2n+1) \to 0$.

Similarly, since

$$\frac{1}{1+t} = 1-t+t^2-\ldots+(-1)^{n-1}t^{n-1}+(-1)^n\frac{t^n}{1+t},$$

therefore

$$\int_0^x \frac{1}{1+t}\,dt = \int_0^x \{1-t+t^2-\ldots+(-1)^{n-1}t^{n-1}\}\,dt+(-1)^n\int_0^x \frac{t^n}{1+t}\,dt,$$

that is,

$$\log(1+x) = x-\frac{x^2}{2}+\frac{x^3}{3}-\ldots+(-1)^{n-1}\frac{x^n}{n}+(-1)^n\int_0^x \frac{t^n}{1+t}\,dt.$$

If $0 \leqslant x \leqslant 1$, then

$$\int_0^x \frac{t^n}{1+t}\,dt \leqslant \int_0^x t^n\,dt, \quad \text{since } \frac{1}{1+t} \leqslant 1 \text{ when } t \geqslant 0,$$

$$= \frac{x^{n+1}}{n+1} \leqslant \frac{1}{n+1}$$

and so $x-\frac{x^2}{2}+\frac{x^3}{3}-\ldots$ converges to $\log(1+x)$, when $0 \leqslant x \leqslant 1$.

Moreover, if $-1 < x \leqslant 0$, then

$$\left|\int_0^x \frac{(-t)^n}{1+t}\,dt\right|$$

$$= \int_0^{-x} \frac{u^n}{1-u}\,du, \quad \text{where } u = -t,$$

$$\leqslant \frac{1}{1+x}\int_0^{-x} u^n\,du, \quad \text{since } 1-u \geqslant 1+x \text{ when } 0 \leqslant u \leqslant -x,$$

$$= \frac{1}{1+x}\frac{(-x)^{n+1}}{n+1} \leqslant \frac{1}{1+x}\frac{1}{1+n},$$

and so, since $1/(1+x)(1+n) \to 0$ for any fixed $x \neq -1$, therefore the series $x-\frac{x^2}{2}+\frac{x^3}{3}-\ldots$ converges to $\log(1+x)$ also when x is negative, provided $x > -1$.

X

DIFFERENTIAL GEOMETRY OF A PLANE CURVE

TANGENT AND NORMAL. AREA AND ARC LENGTH. POLAR COORDINATES. CURVATURE. INFLEXION. EVOLUTES

10. An ordered pair of decimals (x, y) is said to define or to be the coordinates of a *point*, and the point $(0, 0)$ is called the origin (of coordinates). A functional relation $f(x, y) = 0$ is said to define, or to be the equation of, a (plane) *curve*; if X, Y satisfy the equation $f(X, Y) = 0$, the point (X, Y) is said to *lie on* the curve $f(x, y) = 0$, and the curve is said to *pass through* the point. If X, Y satisfy both $f(X, Y) = 0$, $g(X, Y) = 0$ then (X, Y) is called a common point, or point of intersection, of the curves $f(x, y) = 0$ and $g(x, y) = 0$. The curves $f(x, y) = 0$ and $kf(x, y) = 0$ have all their points in common (if $k \neq 0$) and so are said to be the same curve. A pair of equations $x = f(t)$, $y = g(t)$ determine a curve, for the result of eliminating t between the equations is a functional relation between x and y. Such a pair of equations are called the *parametric* equations of a curve.

10.01. The *distance* between two points (x_1, y_1) and (x_2, y_2) is defined to be

$$+\sqrt{\{(x_1-x_2)^2+(y_1-y_2)^2\}}.$$

In particular the distance between (x_1, a) and (x_2, a) is $|x_1-x_2|$ and the distance between (a, y_1) and (a, y_2) is $|y_1-y_2|$.

10.02. If $f(x, y)$ is linear in both x and y, $f(x, y) = 0$ is said to define (or to be the equation of) a *straight line*; since the general linear function is $ax+by+c$, the general equation of a line is $ax+by+c = 0$. If $aa'+bb' = 0$ the lines

$$ax+by+c = 0, \qquad a'x+b'y+c' = 0$$

are said to be *perpendicular*.

The equation of any line through the point (x_1, y_1) is

$$a(x-x_1)+b(y-y_1) = 0$$

since this is the most general linear relation satisfied by x_1, y_1. Of all the lines through a point (x_1, y_1) only one passes through a second point (x_2, y_2), for if $a(x-x_1)+b(y-y_1) = 0$ passes through

(x_2, y_2) then $a(x_2-x_1)+b(y_2-y_1) = 0$ and so the unique line through both points is $\dfrac{x-x_1}{x_2-x_1} = \dfrac{y-y_1}{y_2-y_1}$.

Writing the general equation of a line through the point (x_1, y_1) in the form $\dfrac{x-x_1}{l} = \dfrac{y-y_1}{m}$ we may denote each fraction by r and obtain the equation in *parametric* form

$$x = x_1+lr, \qquad y = y_1+mr.$$

Here x_1, y_1, l, m are constants and r is a variable parameter; the distance of any point (x, y) of the line from the point (x_1, y_1) is

$$+\sqrt{\{(x-x_1)^2+(y-y_1)^2\}} = +\sqrt{\{r^2(l^2+m^2)\}} = |r\sqrt{(l^2+m^2)}|.$$

10.021. We shall use single capital letters as names of points and single small letters as names of lines.

10.03. The distance from a point (X, Y) to a point on the line $x = x_1+lr$, $y = y_1+mr$ (which does not pass through (X, Y)) is

$$h = \sqrt{\{(X-x)^2+(Y-y)^2\}}.$$

Now

$$\frac{dh}{dr} = -\left\{(X-x)\frac{dx}{dr}+(Y-y)\frac{dy}{dr}\right\}\Big/\sqrt{\{(X-x)^2+(Y-y)^2\}}$$
$$= -\{l(X-x)+m(Y-y)\}/\sqrt{\{(X-x)^2+(Y-y)^2\}}$$

and therefore h is a *minimum* for $l(X-x)+m(Y-y) = 0$, i.e.

$$r(l^2+m^2) = l(X-x_1)+m(Y-y_1). \qquad \text{(i)}$$

The line through (X, Y) perpendicular to $(x-x_1)/l = (y-y_1)/m$ is

$$lx+my = lX+mY,$$

and this meets the line $x = x_1+lr$, $y = y_1+mr$ at the point where

$$r(l^2+m^2) = l(X-x_1)+m(Y-y_1),$$

and so by (i) at the point of the line $x = x_1+lr$, $y = y_1+mr$ nearest to (X, Y). Thus the perpendicular from a point P to a line p (i.e. the distance between the point P and the point of intersection of p with the line through P perpendicular to p) is the shortest distance from P to a point on the line p, showing that the use we have made of the term 'perpendicular' is consistent with that in elementary geometry.

10.04. The length of the perpendicular from a point (X, Y) to a line $ax+by+c = 0$ is $|aX+bY+c|/\sqrt{(a^2+b^2)}$.

Let (x_1, y_1) be a point on the line, so that its equation may be expressed in the form $(x-x_1)/l = (y-y_1)/m$, where $m = a$, $l = -b$, and $ly_1-mx_1 = c$. Then by 10.03 the length of the perpendicular is

$$\sqrt{\{(X-x)^2+(Y-y)^2\}},$$

where
$$\frac{Y-y}{X-x} = -\frac{l}{m} = \frac{b}{a},$$

and so
$$\frac{X-x}{a} = \frac{Y-y}{b} = \sigma, \quad \text{say}.$$

Thus the length of the perpendicular is $|\sigma|\sqrt{(a^2+b^2)}$; but

$$ax+by+c = 0,$$

therefore
$$a(X-\sigma a)+b(Y-\sigma b)+c = 0,$$

whence
$$\sigma = (aX+bY+c)/(a^2+b^2),$$

which completes the proof.

10.05. Since the equations $ax+by+c = 0$, $a'x+b'y+c' = 0$ have no common solution when $ab'-a'b = 0$, the lines of which these are the equations have no point of intersection under this condition, and are said to be parallel. Note that each line is perpendicular to the line $bx = ay$.

10.06. The lines $y = 0$, $x = 0$ are called the 'x-axis (of coordinates)' and the 'y-axis (of coordinates)' respectively. The axes are perpendicular.

The perpendicular distances of a point (X, Y) from the axes $x = 0$, $y = 0$ are $|X|$, $|Y|$ respectively.

10.07. The equation of a *circle* centre (a, b) and radius $r > 0$ is

$$(x-a)^2+(y-b)^2 = r^2,$$

for every point on this curve is a distance r from (a, b).

10.1. The straight line

$$\mathbf{x} = \mathbf{f}(\tau)+\mathbf{r}\mathbf{f}'(\tau), \qquad \mathbf{y} = \mathbf{g}(\tau)+\mathbf{r}\mathbf{g}'(\tau)$$

is called the *tangent* to the curve $x = f(t)$, $y = g(t)$ at the point $(f(\tau), g(\tau))$.

For example, the tangent to the parabola $x = t^2$, $y = 2t$ at the point $(9, 6)$ is

$$x = 9+6r, \qquad y = 6+2r,$$

for $dx/dt = 2t$, $dy/dt = 2$, so that when $t = 3$, $dx/dt = 6$, $dy/dt = 2$.

If $f'(\tau)$ and $g'(\tau)$ are both different from zero the equation of the tangent may be written in the form

$$\frac{x-f(\tau)}{f'(\tau)}=\frac{y-g(\tau)}{g'(\tau)},$$

or, denoting $f(\tau)$, $g(\tau)$ by X, Y,

$$y-Y=(x-X)\{g'(\tau)/f'(\tau)\}.$$

If the result of eliminating t between $x=f(t)$, $y=g(t)$ is $y=\phi(x)$, so that $g(t)=\phi\{f(t)\}$ for any t, then $g'(t)=\phi'\{f(t)\}f'(t)$, and so $g'(t)/f'(t)=\phi'(f(t))$, whence $g'(\tau)/f'(\tau)=\phi'(f(\tau))=\phi'(X)$. **Accordingly the tangent to the curve $\mathbf{y}=\boldsymbol{\varphi}(\mathbf{x})$ at $(\mathbf{X},\mathbf{Y})$ is**

$$(\mathbf{y}-\mathbf{Y})=(\mathbf{x}-\mathbf{X})\boldsymbol{\varphi}'(\mathbf{X}).$$

EXAMPLE. The equation of the tangent to the parabola $y=4x^2$ at the point (3, 36) is

$$y-36=(x-3)8.3,$$

i.e.

$$24x-y-36=0.$$

When $\phi'(x)$ exists, it is called the *slope* or gradient of the curve $y=\phi(x)$ at the point (x,y). Thus if $b\neq 0$ the slope of the line $ax+by+c=0$ is $-a/b$.

Even when the equation of the curve $f(x,y)=0$ cannot be solved for y, we may be able to find the tangent at a point of the curve; for instance, if $x^5+y^5=33$ then $5x^4+5y^4(dy/dx)=0$ and therefore at the point (2, 1) on the curve the gradient dy/dx has the value $-2^4/1^4=-16$ so that the equation of the tangent is

$$(y-1)=-16(x-2),\quad \text{i.e.}\quad 16x+y-33=0.$$

10.2. The *acute angle* between the lines $ax+by+c=0$ and $a'x+b'y+c'=0$ is defined to be

$$\cos^{-1}|(aa'+bb')/\sqrt{\{(a^2+b^2)(a'^2+b'^2)\}}|$$

and the *obtuse angle* between the lines is

$$\cos^{-1}[-|(aa'+bb')/\sqrt{\{(a^2+b^2)(a'^2+b'^2)\}}|].$$

Since $\cos^{-1}(-x)=\pi-\cos^{-1}x$, it follows that the sum of the acute and obtuse angles between any two lines is π; if the lines are perpendicular, so that $aa'+bb'=0$, then the angles between the lines are both equal to $\frac{1}{2}\pi$—conversely, if an angle between two lines is $\frac{1}{2}\pi$ then $|(aa'+bb')/\sqrt{\{(a^2+b^2)(a'^2+b'^2)\}}|=\cos\frac{1}{2}\pi=0$ and

therefore $aa'+bb' = 0$ which proves that the lines are perpendicular. Observe that since $\cos\theta \geqslant 0$ if $0 \leqslant \theta \leqslant \frac{1}{2}\pi$, and $\cos\theta \leqslant 0$ if $\frac{1}{2}\pi \leqslant \theta \leqslant \pi$, therefore an acute angle lies between 0 and $\frac{1}{2}\pi$ and an obtuse angle between $\frac{1}{2}\pi$ and π.

10.21. The acute angle between the line $ax+by+c = 0$ and the x-axis, $y = 0$, is $\cos^{-1}|b/\sqrt{(a^2+b^2)}|$; if $\theta = \cos^{-1}|b\sqrt{(a^2+b^2)}|$, so that $0 \leqslant \theta \leqslant \frac{1}{2}\pi$, then $\cos\theta = |b|/\sqrt{(a^2+b^2)}$, and so (provided $b \neq 0$)

$$\tan^2\theta = \sec^2\theta - 1 = \{(a^2+b^2)/b^2\} - 1 = a^2/b^2,$$

whence (since $0 \leqslant \theta \leqslant \frac{1}{2}\pi$) $\tan\theta = |a/b|$ and $\tan(-\theta) = -|a/b|$; if $b = 0$, $\theta = \frac{1}{2}\pi$. The slope of the line $ax+by+c = 0$ is $-a/b$ ($b \neq 0$), and therefore, if θ is the acute angle between $ax+by+c = 0$ and the x-axis, the slope of $ax+by+c = 0$ is $\tan\theta$ if a/b is negative and is $\tan(-\theta)$ if a/b is positive. Hence if we define ψ by the conditions

$$\begin{aligned} \psi &= \theta && \text{if } a/b \text{ is negative or zero,} \\ \psi &= -\theta && \text{if } a/b \text{ is positive,} \\ \psi &= \tfrac{1}{2}\pi && \text{if } b = 0, \end{aligned}$$

then the slope of the line $ax+by+c = 0$ is $\tan\psi$ (if $\psi < \frac{1}{2}\pi$); ψ is called the *inclination* of the line. Observe that $-\frac{1}{2}\pi < \psi \leqslant \frac{1}{2}\pi$. It follows that the equation of a line of inclination ψ is

$$y = x\tan\psi + c$$

if $\psi < \frac{1}{2}\pi$, and is $ax+c = 0$ if $\psi = \frac{1}{2}\pi$.

10.22.* A transformation from the variables x, y to new variables X, Y by the relations

$$X = x+a, \qquad Y = y+b,$$

for any constants a, b, is called a *change of origin*.

A transformation by the relations

$$X = x\cos\alpha + y\sin\alpha, \qquad Y = x\sin\alpha - y\cos\alpha,$$

for any constant α, is called a *rotation of axes*.

A combination of change of origin and a rotation of axes, expressed by the relations

$$X = x\cos\alpha + y\sin\alpha + a, \qquad Y = x\sin\alpha - y\cos\alpha + b,$$

is called a general *transformation of rectangular axes*.

Notice that a general transformation may be effected by a rotation followed by a change of origin, or vice versa. For if $\xi = x\cos\alpha + y\sin\alpha$, $\eta = x\sin\alpha - y\cos\alpha$ and $X = \xi + a$, $Y = \eta + b$, then

$$X = x\cos\alpha + y\sin\alpha + a, \qquad Y = x\sin\alpha - y\cos\alpha + b;$$

and if $\xi = x + (a\cos\alpha + b\sin\alpha)$, $\eta = y + (a\sin\alpha - b\cos\alpha)$ and $X = \xi\cos\alpha + \eta\sin\alpha$, $Y = \xi\sin\alpha - \eta\cos\alpha$, then

$$X = x\cos\alpha + y\sin\alpha + a, \qquad Y = x\sin\alpha - y\cos\alpha + b.$$

10.23.* The distance between two points, and the angle between two lines, are unchanged by a transformation of axes.

For if $X_1 = x_1 + a$, $X_2 = x_2 + a$ and $Y_1 = y_1 + b$, $Y_2 = y_2 + b$ then

$$(X_1 - X_2)^2 + (Y_1 - Y_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2,$$

and if $X_1 = x_1\cos\alpha + y_1\sin\alpha$, $X_2 = x_2\cos\alpha + y_2\sin\alpha$, etc., then

$$\begin{aligned}(X_1 - X_2)^2 &+ (Y_1 - Y_2)^2 \\ &= \{(x_1 - x_2)\cos\alpha + (y_1 - y_2)\sin\alpha\}^2 + \{(x_1 - x_2)\sin\alpha - (y_1 - y_2)\cos\alpha\}^2 \\ &= (x_1 - x_2)^2 + (y_1 - y_2)^2.\end{aligned}$$

The pair of equations $ax + by + c = 0$ and $a'x + b'y + c' = 0$ transform under $x = X\cos\alpha + Y\sin\alpha + l$, $y = X\sin\alpha - Y\cos\alpha + m$ into

$$X(a\cos\alpha + b\sin\alpha) + Y(a\sin\alpha - b\cos\alpha) + al + bm + c$$

and

$$X(a'\cos\alpha + b'\sin\alpha) + Y(a'\sin\alpha - b'\cos\alpha) + a'l + b'm + c'$$

and since

$$\begin{aligned}(a\cos\alpha + b\sin\alpha)(a'\cos\alpha + b'\sin\alpha) &+ \\ + (a\sin\alpha - b\cos\alpha)(a'\sin\alpha - b'\cos\alpha) &= aa' + bb'\end{aligned}$$

and

$$(a\cos\alpha + b\sin\alpha)^2 + (a\sin\alpha - b\cos\alpha)^2 = a^2 + b^2,$$

$$(a'\cos\alpha + b'\sin\alpha)^2 + (a'\sin\alpha - b'\cos\alpha)^2 = a'^2 + b'^2,$$

it follows that the (acute) angle between the lines is unchanged.

10.24.* Let (x_r, y_r) be called the point P_r, $r = 1, 2, 3$, and denote the distance between P_r and P_s by P_rP_s and the line joining P_r to P_s by (P_r, P_s). We prove:

If (P_1, P_2) is perpendicular to (P_1, P_3) and if θ is the acute angle between (P_1, P_2) and (P_2, P_3) then

$$P_2P_3^2 = P_1P_2^2+P_1P_3^2$$

and $$P_1P_2 = P_2P_3\cos\theta, \qquad P_1P_3 = P_2P_3\sin\theta.$$

Since (P_1, P_2) is perpendicular to (P_1, P_3)

$$\frac{x_3-x_1}{y_2-y_1} = -\frac{y_3-y_1}{x_2-x_1} = \frac{b}{a}, \quad \text{say.} \tag{i}$$

Let (a, α) be the polar coordinates of the point (x_2-x_1, y_2-y_1) so that

$$x_2-x_1 = a\cos\alpha, \qquad y_2-y_1 = a\sin\alpha$$

(see § 10.5); then by equations (i)

$$x_3-x_1 = b\sin\alpha, \qquad y_3-y_1 = -b\cos\alpha.$$

Consider the transformation

$$X = (x-x_1)\cos\alpha+(y-y_1)\sin\alpha,$$
$$Y = (x-x_1)\sin\alpha-(y-y_1)\cos\alpha.$$

The coordinates

$$x = x_1, x_2, x_3, \qquad y = y_1, y_2, y_3$$

take the values

$$X = 0,\ a,\ 0, \qquad Y = 0,\ 0,\ b \quad \text{respectively.}$$

The equation of the line joining the points $(a, 0)$, $(0, b)$ is

$$bX+aY = ab,$$

and the acute angle θ between this line and the line $Y = 0$ is given by $\cos\theta = |a|/\sqrt{(a^2+b^2)}$. The distance between $(a, 0)$ and $(0, b)$ is $\sqrt{(a^2+b^2)}$ and so, by 10.23,

$$P_2P_3^2 = a^2+b^2 = P_1P_2^2+P_1P_3^2$$

and

$$P_1P_2 = |a| = P_2P_3\cos\theta, \qquad P_1P_3 = |b| = P_2P_3\sin\theta.$$

10.3. The area bounded by a curve

If $x(t)$, $y(t)$ are differentiable functions, and if $x(t_0) = x(t_1)$, $y(t_0) = y(t_1)$, but the equations $x(t) = x(u)$, $y(t) = y(u)$ cannot be simultaneously satisfied by a t and u in the interval $(t_0, t_1]$, open to the right, then the equations $x = x(t)$, $y = y(t)$, $t_0 \leqslant t \leqslant t_1$, define a *simple closed curve*.

The formula for the *area* bounded by a simple closed curve is

$$\text{(A)} \qquad \left|\tfrac{1}{2}\int_{t_0}^{t_1}(\mathbf{x}\mathbf{y}'-\mathbf{x}'\mathbf{y})\,\mathbf{dt}\right|,$$

where $x' = dx/dt$, $y' = dy/dt$.

A rigorous proof of this formula belongs to the theory of double integrals and is postponed to Chapter XVII.

It is sufficient for the application of formula (A) if $x(t)$ and $y(t)$ are continuous in the interval (t_0, t_1) and if the interval may be divided into a finite number of parts in each of which the functions $x(t)$ and $y(t)$ are differentiable.

For instance the formula may be applied to the functions $x(t)$, $y(t)$ defined in (0, 4) as follows:

$$\begin{aligned}
x(t) &= a+(b-a)t, & y(t) &= c, & 0 &\leqslant t \leqslant 1,\\
&= b, & &= c+(d-c)(t-1), & 1 &\leqslant t \leqslant 2,\\
&= b-(b-a)(t-2), & &= d, & 2 &\leqslant t \leqslant 3,\\
&= a, & &= d-(d-c)(t-3), & 3 &\leqslant t \leqslant 4.
\end{aligned}$$

Both $x(t)$ and $y(t)$ are differentiable in each of the intervals (0, 1), (1, 2), (2, 3), and (3, 4), and are continuous in (0, 4), but not differentiable in (0, 4). The curve described by $x(t)$, $y(t)$ consists of the parallel segments $y = c$, $y = d$, $a \leqslant x \leqslant b$, and the parallel segments $x = a$, $x = b$, $c \leqslant y \leqslant d$, which, since the lines $x = a$, $y = c$ are perpendicular, *determine a rectangle.*

Now

$$\begin{aligned}
&\int_0^4 (xy'-x'y)\,dt\\
&= \int_0^1 + \int_1^2 + \int_2^3 + \int_3^4 (xy'-x'y)\,dt\\
&= -\int_0^1 c(b-a)\,dt + \int_1^2 b(d-c)\,dt + \int_2^3 d(b-a)\,dt - \int_3^4 a(d-c)\,dt\\
&= 2(b-a)(d-c).
\end{aligned}$$

Hence formula (A) gives the familiar value $|(b-a)(d-c)|$ for the area of the rectangle.

Example. To find the area bounded by a loop of the curve

$$x = a\sin 2t, \qquad y = a\sin t.$$

Since x and y are zero simultaneously when $t = 0$ and $t = \pi$ but not for any value of t between 0 and π, therefore the equations $x = a\sin 2t$, $y = a\sin t$, $0 \leqslant t \leqslant \pi$, determine a closed curve which bounds an area

$$\tfrac{1}{2}\int_0^{\pi} (xy'-x'y)\, dt$$

$$= \tfrac{1}{2}a^2 \int_0^{\pi} (\sin 2t \cos t - 2\cos 2t \sin t)\, dt$$

$$= a^2 \int_0^{\pi} \sin t\, dt - a^2 \int_0^{\pi} \sin t \cos^2 t\, dt,$$

$$\text{since } 2(\sin 2t \cos t - \cos 2t \sin t) = 2\sin t,$$

$$= a^2[-\cos t]_0^{\pi} + a^2\left[\frac{\cos^3 t}{3}\right]_0^{\pi} = 2a^2 \quad a^2 = \tfrac{4}{3}a^2.$$

10.31. If $x = x(t)$, $y = tx$, $t_0 \leqslant t \leqslant t_1$, are the equations of a closed curve, then the area bounded by the curve is the positive value of the integral

$$\tfrac{1}{2}\int_{t_0}^{t_1} x^2\, dt.$$

For
$$xy'-x'y = x^2(d/dt)(y/x) = x^2, \quad \text{since } y/x = t,$$

and therefore

$$\tfrac{1}{2}\int_{t_0}^{t_1} x^2\, dt = \tfrac{1}{2}\int_{t_0}^{t_1} (xy'-x'y)\, dt,$$

which completes the proof.

Example. To find the area of the loop of *Descartes' Folium*

$$x^3+y^3 = 3axy.$$

Write $y = tx$, then $x^3(1+t^3) = 3ax^2t$ so that

$$x = 3at/(1+t^3), \qquad y = 3at^2/(1+t^3).$$

When $t = 0$, $x = y = 0$, and $\lim \dfrac{3an}{1+n^3} = \lim \dfrac{3an^2}{1+n^3} = 0$. Furthermore x and y are not simultaneously zero for any positive t, and so $x = 3at/(1+t^3)$, $y = 3at^2/(1+t^3)$ describes a closed curve as t

increases from 0 through all positive values, and the area bounded is

$$\lim \frac{1}{2}\int_0^n \frac{9a^2t^2}{(1+t^3)^2}\,dt = \lim\left[\frac{-3a^2}{2(1+t^3)}\right]_0^n = \frac{3a^2}{2}.$$

10.32. *The area bounded by the curves $y = f(x)$, $y = g(x)$ and the lines $x = a$, $x = b$.*

We suppose that $f(x) > g(x)$ in $[a, b]$. To obtain a parametric representation of the curve bounding the given area, we define

$$\begin{aligned}
x(t) &= a+(b-a)t, & y(t) &= g\{x(t)\}, & 0 &\leqslant t \leqslant 1,\\
&= b, & &= g(b)+\{f(b)-g(b)\}(t-1), & 1 &\leqslant t \leqslant 2,\\
&= b-(b-a)(t-2), & &= f\{x(t)\}, & 2 &\leqslant t \leqslant 3,\\
&= a, & &= f(a)+\{g(a)-f(a)\}(t-3), & 3 &\leqslant t \leqslant 4.
\end{aligned}$$

As t varies from 0 to 1, the point $x(t)$, $y(t)$ describes the curve $y = g(x)$, x varying from a to b; as t varies from 1 to 2, $x(t)$, $y(t)$ describes the line $x = b$, y varying from $g(b)$ to $f(b)$; as t varies from 2 to 3, $x(t)$, $y(t)$ describes the curve $y = f(x)$, x varying from b to a, and, finally, as t varies from 3 to 4, $x(t)$, $y(t)$ describes the line $x = a$, y varying from $f(a)$ to $g(a)$.

Thus we have determined a parametric representation of the given curve. The area bounded by the curve is the positive value of the integral

$$\begin{aligned}
&\tfrac{1}{2}\int_0^4 (xy'-x'y)\,dt\\
&\quad= \tfrac{1}{2}\Big[\int_0^1 \{xg'(x)x'-x'g(x)\}\,dt + \int_1^2 b\{f(b)-g(b)\}\,dt +\\
&\qquad + \int_2^3 \{xf'(x)x'-x'f(x)\}\,dt + \int_3^4 a\{g(a)-f(a)\}\,dt\Big].
\end{aligned}$$

In the first and third integrals we transform from the variable t to the variable x by the relation $x = x(t)$, and we obtain

$$\begin{aligned}
\int_0^1 \{xg'(x)-g(x)\}x'\,dt &= \int_a^b \{xg'(x)-g(x)\}\,dx\\
&= [xg(x)]_a^b - 2\int_a^b g(x)\,dx, \quad \text{integrating by parts,}
\end{aligned}$$

and

$$\int_{2}^{3} \{xf'(x)-f(x)\}x'\,dt = \int_{b}^{a} \{xf'(x)-f(x)\}\,dx$$
$$= -[xf(x)]_a^b + 2\int_a^b f(x)\,dx.$$

Furthermore

$$\int_{}^{1} b\{f(b)-g(b)\}\,dt + \int_{}^{} a\{g(a)-f(a)\}\,dt = [xf(x)]_a^b - [xg(x)]_a^b,$$

and therefore the area is given by

10.33. $$\int_a^b \{\mathbf{f(x)-g(x)}\}\,\mathbf{dx}.$$

10.34. An important special case of 10.33 arises when the curve $y = g(x)$ is the x-axis, $y = 0$. In this case it follows from 10.33 that the area bounded by a curve $y = f(x)$, the x-axis, and the lines $x = a$, $x = b$ is

$$\int_a^b f(x)\,dx$$

provided $f(x) > 0$ in $[a, b]$.

Similarly, taking $f(x) = 0$, $g(x) < 0$, we find the area bounded by the x-axis, the lines $x = a$, $x = b$, and the curve $y = g(x)$ is

$$\left|\int_a^b g(x)\,dx\right| = -\int_a^b g(x)\,dx.$$

Examples. I. The area bounded by the parabola $y = x^2$, the x-axis, and the lines $x = 0$, $x = c$ is

$$\int_0^c x^2\,dx = \left[\frac{x^3}{3}\right]_0^c = \tfrac{1}{3}c^3,$$

which is one-third of the area of the rectangle bounded by the lines $x = 0$, $x = c$, the x-axis, and the line parallel to the x-axis through the point (c, c^2) on the parabola.

II. The area of the circle $x^2+y^2 = a^2$.

The circle is bounded by the two curves

$$y = \sqrt{(a^2-x^2)}, \qquad y = -\sqrt{(a^2-x^2)}$$

which intersect at (and only at) the points $(a, 0)$, $(-a, 0)$. The area of the circle is therefore

$$\int_{-a}^{a} [\sqrt{(a^2-x^2)}-\{-\sqrt{(a^2-x^2)}\}]\, dx$$

$$= 2\int_{-a}^{a} \sqrt{(a^2-x^2)}\, dx = [x\sqrt{(a^2-x^2)}+a^2 \arcsin x/a]_{-a}^{a} = \pi a^2.$$

10.4. The formula for *the length of the arc of the curve* $x = x(t)$, $y = y(t)$ from $t = t_0$ to t is

$$\text{(L)} \qquad s(t) = \int_{t_0}^{t} +\sqrt{(x'^2+y'^2)}\, dt.$$

The proof of this formula is postponed to Chapter XV. It is sufficient for the application of this formula that the interval (t_0, t) can be divided into a finite number of parts in each of which $x(t)$, $y(t)$ are differentiable.

10.41. Since a parametric representation of the arc of the curve $y = f(x)$, $a \leqslant x \leqslant b$, is given by

$$x(t) = t, \qquad y(t) = f(t), \qquad a \leqslant t \leqslant b,$$

it follows that the length of the arc is

$$\int_{a}^{b} +\sqrt{(x'^2+y'^2)}\, dt = \int_{a}^{b} +\sqrt{[1+\{f'(t)\}^2]}\, dt$$

$$\textbf{10.411.} \qquad = \int_{a}^{b} +\sqrt{\left\{1+\left(\frac{dy}{dx}\right)^2\right\}}\, dx.$$

Conversely we may derive formula (L) from formula 10.411. For if $x = x(t)$, $y = y(t)$ and $a = x(t_0)$, $b = x(t_1)$, then since $dy/dx = y'/x'$ we have

$$\int_{a}^{b} \sqrt{\left\{1+\left(\frac{dy}{dx}\right)^2\right\}}\, dx = \int_{t_0}^{t_1} \sqrt{\left\{1+\left(\frac{y'}{x'}\right)^2\right\}}x'\, dt = \int_{t_0}^{t_1} \sqrt{(x'^2+y'^2)}\, dt.$$

EXAMPLES. I. The length of the arc of the parabola $y = x^2$ from $x = 0$ to $x = c$ is

$$\int_0^c \sqrt{\left\{1+\left(\frac{dy}{dx}\right)^2\right\}}\,dx = \int_0^c \sqrt{(1+4x^2)}\,dx$$

$$= \tfrac{1}{2}\int_0^{2c} \sqrt{(1+t^2)}\,dt, \quad x = \tfrac{1}{2}t,$$

$$= \tfrac{1}{4}[t\sqrt{(1+t^2)}+\log\{t+\sqrt{(1+t^2)}\}]_0^{2c}$$

$$= \tfrac{1}{2}c\sqrt{(1+4c^2)}+\tfrac{1}{4}\log\{2c+\sqrt{(1+4c^2)}\}.$$

II. The length of the circumference of the circle $x^2+y^2 = a^2$.

The circumference of the circle is made up of two semicircular arcs, $y = +\sqrt{(a^2-x^2)}$, $y = -\sqrt{(a^2-x^2)}$, each from $x = -a$ to $x = +a$. On the first semicircle $dy/dx = -x/\sqrt{(a^2-x^2)}$ and so the length of the arc is

$$\int_{-a}^{a} \sqrt{[1+\{x^2/(a^2-x^2)\}]}\,dx = a\int_{-a}^{a} \frac{dx}{\sqrt{(a^2-x^2)}} = a[\arcsin(x/a)]_{-a}^{a} = \pi a.$$

Similarly the length of the second arc is πa, and so the total circumference is $2\pi a$.

Alternatively, since $x = a\cos t$, $y = a\sin t$, $0 \leqslant t \leqslant 2\pi$, is a parametric representation of the circle, the length of the circumference is given by

$$\int_0^{2\pi} +\sqrt{(a^2\sin^2 t+a^2\cos^2 t)}\,dt = a\int_0^{2\pi} dt = 2\pi a.$$

THIS COMPLETES THE PROOF THAT THE DECIMAL π WE INTRODUCED IN § 5.2, AS THE LEAST POSITIVE ROOT OF THE EQUATION $\cos\frac{1}{2}x = 0$, IS THE SAME AS THE NUMBER π IN ELEMENTARY GEOMETRY, NAMELY, THE 'RATIO' OF THE CIRCUMFERENCE OF A CIRCLE TO ITS DIAMETER.

10.42. *In a unit circle a chord of length* $2\sin x$ *cuts off an arc of length* $2x$ (provided $0 \leqslant x \leqslant \frac{1}{2}\pi$).

Consider the chord joining the points $(\sin\alpha, \cos\alpha)$, $(-\sin\alpha, \cos\alpha)$, $0 \leqslant \alpha \leqslant \frac{1}{2}\pi$, which lie on the unit circle $x^2+y^2 = 1$.

The length of the chord joining these points is $2\sin\alpha$, and the length of the arc joining them is

$$\int\limits_{-\sin\alpha}^{\sin\alpha} \sqrt{\{1+(dy/dx)^2\}}\,dx = \int\limits_{-\sin\alpha}^{\sin\alpha} \sqrt{\{1+x^2/(1-x^2)\}}\,dx$$

$$= \int\limits_{-\sin\alpha}^{\sin\alpha} \{1/\sqrt{(1-x^2)}\}\,dx = [\arcsin x]_{-\sin\alpha}^{\sin\alpha} = 2\alpha$$

THIS COMPLETES THE IDENTIFICATION OF THE FUNCTIONS $\sin x$ AND $\cos x$, INTRODUCED IN § 5, WITH THE ELEMENTARY TRIGONOMETRIC FUNCTIONS.

EXAMPLE. To find the area bounded by the x-axis and an arch of the cycloid $x = a(t-\sin t)$, $y = a(1-\cos t)$.

The cycloid meets the x-axis at the points $t = 2n\pi$, so that a complete arch is contained between the points $x = 0$, $x = 2a\pi$ where $t = 0$, $t = 2\pi$ respectively. The required area is

$$\int^{2a\pi} y\,dx = \int^{2\pi} yx'\,dt = \int\limits_0^{2\pi} a^2(1-\cos t)^2\,dt$$

$$= a^2 \int\limits_0^{2\pi} (1-2\cos t+\cos^2 t)\,dt$$

$$= \frac{a^2}{2}\int\limits_0^{2\pi} (3-4\cos t+\cos 2t)\,dt$$

$$= \frac{a^2}{2}[3t-4\sin t+\tfrac{1}{2}\sin 2t]_0^{2\pi}$$

$$= 3\pi a^2.$$

The length of the arch of the cycloid is

$$\int^{2\pi} \sqrt{(x'^2+y'^2)}\,dt = a\int^{2\pi} \{(1-\cos t)^2+\sin^2 t\}^{\frac{1}{2}}\,dt$$

$$= a\int\limits_0^{2\pi} \{2(1-\cos t)\}^{\frac{1}{2}}\,dt$$

$$= 2a\int\limits_0^{2\pi} \sin\tfrac{1}{2}t\,dt, \quad \text{since } \sin\tfrac{1}{2}t \text{ is positive in } (0, 2\pi),$$

$$= 4a[-\cos\tfrac{1}{2}t]_0^{2\pi} = 8a.$$

10.43. The *angle subtended at the centre* of the circle

$$x = a+R\cos t, \qquad y = b+R\sin t \qquad (0 \leqslant t < 2\pi)$$

by the arc of the circle from the point $t = \alpha$ to the point $t = \beta$, $\beta > \alpha$, is defined to be $\beta-\alpha$.

The acute angle between the lines joining these points to the centre of the circle is

$$\arccos|(1+\tan\alpha\tan\beta)/\sqrt{\{(1+\tan^2\alpha)(1+\tan^2\beta)\}}|$$
$$= \arccos|\cos(\beta-\alpha)| = \gamma, \text{ say.}$$

Thus the relation between γ and $\beta-\alpha$ is given by the equations

$$\gamma = \beta-\alpha \qquad \text{if} \quad 0 \leqslant \beta-\alpha \leqslant \tfrac{1}{2}\pi,$$
$$\gamma = \pi-(\beta-\alpha) \qquad \text{if} \quad \tfrac{1}{2}\pi < \beta-\alpha \leqslant \pi,$$
$$\gamma = (\beta-\alpha)-\pi \qquad \text{if} \quad \pi < \beta-\alpha \leqslant \tfrac{3}{2}\pi,$$

and
$$\gamma = 2\pi-(\beta-\alpha) \qquad \text{if} \quad \tfrac{3}{2}\pi < \beta-\alpha < 2\pi.$$

10.44. The length of the arc of the circle

$$x = a+R\cos t, \qquad y = b+R\sin t$$

from $t = \alpha$ to $t = \alpha+\theta$ is

$$\int_{\alpha}^{\alpha+\theta} \sqrt{(x'^2+y'^2)}\,dt = R\int_{\alpha}^{\alpha+\theta} \sqrt{(\sin^2 t+\cos^2 t)}\,dt = R\int_{\alpha}^{\alpha+\theta} dt = R\theta.$$

By 10.43, the angle subtended by this arc at the centre of the circle is θ, and therefore

THE LENGTH OF THE ARC OF A CIRCLE IS THE PRODUCT OF THE RADIUS OF THE CIRCLE AND THE ANGLE SUBTENDED BY THE ARC AT THE CENTRE.

10.45. Units. Elementary trigonometry introduces two units for recording the size of angles. One unit is the *degree*, which is a ninetieth part of the angle between two perpendicular lines, and the other is the *radian*, which is the angle subtended at the centre of a circle by an arc equal in length to the radius. The relation between these units is that an angle of π radians is equal to an angle of 180 degrees.

We have seen in § 10.44 that in a circle of radius R, an arc of length R subtends a unit angle at the centre. This shows that **the unit angle of the present work is equal to the radian of elementary trigonometry.**

10.5. Polar coordinates

The pair of numbers r, θ defined in terms of x and y by the relations

$$\begin{aligned} r &= +\sqrt{(x^2+y^2)}, \\ \theta &= \arccos(x/r), && \text{if } y \geqslant 0,\ r > 0, \\ &= 2\pi - \arccos(x/r), && \text{if } y < 0, \end{aligned}$$

are called the *polar coordinates* of the point (x, y). x and y are expressed in terms of r, θ by the equations

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

The equation $\phi(r, \theta) = 0$, where $\phi(r, \theta) = f(r\cos\theta, r\sin\theta)$, is called the polar equation of the curve $f(x, y) = 0$.

The general polar equation of a straight line is

$$r(A\cos\theta + B\sin\theta) + C = 0;$$

in particular lines through the origin have the equation

$$\theta = \text{constant},$$

for if $\theta = \alpha$ then $y/x = \tan\theta = \tan\alpha$, i.e. $y = x\tan\alpha$.

10.51. To find the area bounded by the curve $r = f(\theta)$ and the lines $\theta = \alpha$, $\theta = \beta$.

A parametric representation of the bounding curve is given by

$$x(t) = r(t)\cos\theta(t), \qquad y(t) = r(t)\sin\theta(t),$$

where

$$\begin{aligned} \theta(t) &= \alpha, & r(t) &= tf(\alpha), & 0 \leqslant t \leqslant 1, \\ &= \alpha + (\beta-\alpha)(t-1), & &= f\{\theta(t)\}, & 1 \leqslant t \leqslant 2, \\ &= \beta, & &= (3-t)f(\beta), & 2 \leqslant t \leqslant 3. \end{aligned}$$

The required area is therefore

$$\tfrac{1}{2}\int_0^3 (xy' - x'y)\,dt = \tfrac{1}{2}\int_0^3 x^2 \frac{d}{dt}(y/x)\,dt = \tfrac{1}{2}\int_0^3 r^2\cos^2\theta\,\frac{d}{dt}(\tan\theta)\,dt$$

$$= \tfrac{1}{2}\int_1^2 r^2\cos^2\theta\sec^2\theta\,\frac{d\theta}{dt}\,dt,$$

since $\theta(t)$ is constant in $(0, 1)$ and in $(2, 3)$

$$= \tfrac{1}{2}\int_1^2 r^2\frac{d\theta}{dt}\,dt = \tfrac{1}{2}\int_\alpha^\beta r^2\,d\theta, \quad \text{since } \theta(1) = \alpha,\ \theta(2) = \beta.$$

Thus the area bounded by the curve $r = f(\theta)$ and the lines $\theta = \alpha$, $\theta = \beta$ is

$$\tfrac{1}{2}\int\limits_{\alpha}^{\beta} \mathbf{r}^2\, d\boldsymbol{\theta}.$$

10.52. To find the length of the arc of the curve $r = f(\theta)$ from $\theta = \alpha$ to $\theta = \beta$.

A parametric representation of the arc, with parameter θ, is

$$x(\theta) = f(\theta)\cos\theta, \qquad y(\theta) = f(\theta)\sin\theta.$$

Since

$$s'^2 = x'^2+y'^2 = (r'\cos\theta - r\sin\theta)^2+(r'\sin\theta+r\cos\theta)^2 = r'^2+r^2,$$

the dash denoting differentiation with respect to θ, therefore the length of the arc is

$$\int\limits_{\alpha}^{\beta} +\sqrt{(\mathbf{r}'^2+\mathbf{r}^2)}\, d\boldsymbol{\theta}.$$

EXAMPLES. I. To find the areas of the loops of the curve

$$r^2 = (1+2\cos\theta)^2.$$

The expression $1+2\cos\theta$ is negative when $\frac{2}{3}\pi \leqslant \theta \leqslant \frac{4}{3}\pi$ and positive when θ lies outside this range, that is when $0 \leqslant \theta \leqslant \frac{2}{3}\pi$ or $\frac{4}{3}\pi \leqslant \theta \leqslant 2\pi$. Hence the loops are given by

$$r = 1+2\cos\theta, \quad 0 \leqslant \theta \leqslant \tfrac{2}{3}\pi \quad \text{or} \quad \tfrac{4}{3}\pi \leqslant \theta \leqslant 2\pi,$$

and
$$r = -2\cos\theta - 1, \quad \tfrac{2}{3}\pi \leqslant \theta \leqslant \tfrac{4}{3}\pi.$$

The area of the first loop is therefore

$$\tfrac{1}{2}\int\limits^{\frac{2}{3}\pi} r^2\, d\theta + \tfrac{1}{2}\int\limits_{\frac{4}{3}\pi}^{2\pi} r^2\, d\theta$$

$$= \tfrac{1}{2}\Big\{\int\limits_{0}^{\frac{2}{3}\pi} + \int\limits_{\frac{4}{3}\pi}^{2\pi}(1+4\cos\theta+2(1+\cos 2\theta))\, d\theta$$

$$= \tfrac{1}{2}\{2\pi+4\sin\tfrac{2}{3}\pi+\sin\tfrac{4}{3}\pi+2\pi-4\sin\tfrac{4}{3}\pi-\sin\tfrac{8}{3}\pi\}$$

$$= 2\pi+4\sin\tfrac{1}{3}\pi-\sin\tfrac{1}{3}\pi = 2\pi+\frac{3\sqrt{3}}{2}.$$

And the area of the second loop is

$$\tfrac{1}{2}\int\limits_{\frac{2}{3}\pi}^{\frac{4}{3}\pi} r^2\, d\theta = \tfrac{1}{2}\{2\pi+4\sin\tfrac{4}{3}\pi+\sin\tfrac{8}{3}\pi-4\sin\tfrac{2}{3}\pi-\sin\tfrac{4}{3}\pi\}$$

$$= \pi - \frac{3\sqrt{3}}{}$$

II. To find the total length of the cardioid

$$r = 1+\cos\theta, \quad 0 \leqslant \theta \leqslant 2\pi.$$

We have

$$s'^2 = r^2+r'^2 = (1+\cos\theta)^2+\sin^2\theta$$
$$= 2(1+\cos\theta) = 4\cos^2\tfrac{1}{2}\theta.$$

Hence the required length is

$$2\int_0^{2\pi} +\sqrt{(\cos^2\tfrac{1}{2}\theta)}\,d\theta = 2\left\{\int_0^{\pi} \cos\tfrac{1}{2}\theta\,d\theta + \int_{\pi}^{2\pi} -\cos\tfrac{1}{2}\theta\,d\theta\right\},$$

since $\cos\frac{1}{2}\theta$ is positive in the range $0 \leqslant \theta \leqslant \pi$ and negative in the range $\pi \leqslant \theta \leqslant 2\pi$,

$$= 4\{[\sin\tfrac{1}{2}\theta]_0^{\pi} - [\sin\tfrac{1}{2}\theta]_{\pi}^{2\pi}\}$$
$$= 8.$$

10.6. A differentiable relation between three variables,

$$f(x, y, z) = 0,$$

is called a *surface*. If the equation has one of the forms

$$y^2+z^2 = f(x)^2, \qquad x^2+z^2 = f(y)^2, \qquad x^2+y^2 = f(z)^2$$

the surface is known as a *surface of revolution* with axial symmetry. The surface $y^2+z^2 = f(x)^2$ is said to be formed by *rotating* the curve $y = f(x)$ (or the curve $z = f(x)$) *about the x-axis* (and similarly for the other forms in relation to the y- and z-axes).

10.61. The formula for the *volume* contained by the surface formed by rotating the arc of the curve $y = f(x)$, from $x = a$ to $x = b$, about the x-axis is

$$\int_a^b \pi y^2\,dx$$

and the area of the surface generated by the rotating curve is given by

$$\int_{s_a}^{s_b} 2\pi y\,ds,$$

where s is the arc length of the curve $y = f(x)$ and s_a, s_b the distances along the curve from the point where $x = 0$ to the points where $x = a$, $x = b$ respectively; since $\dfrac{ds}{dx} = \sqrt{\left\{1+\left(\dfrac{dy}{dx}\right)^2\right\}}$,

this integral, after transforming the variable from s to x, becomes

$$\int^{b} 2\pi y\sqrt{\left\{1+\left(\frac{dy}{dx}\right)^2\right\}}\,dx.$$

The foregoing formulae for volume and surface area are provable by multiple integrals.

EXAMPLE. The surface formed by rotating the arc $y = \sqrt{(a^2-x^2)}$, from $x = -a$ to $x = a$, about the x-axis, has the equation $y^2+z^2 = a^2-x^2$, that is, $x^2+y^2+z^2 = a^2$, which is the equation of a sphere. Thus the volume of the zone of a sphere, formed by rotating the arc $y = \sqrt{(a^2-x^2)}$ from $x = p_1$ to $x = p_2$, about the x-axis is

$$\int_{p_1}^{p_2} \pi(a^2-x^2)\,dx = \tfrac{1}{3}\pi(p_2-p_1)\{3a^2-(p_1^2+p_1p_2+p_2^2)\}$$

and the surface area of this zone is

$$\int_{p_1}^{p_2} 2\pi\sqrt{(a^2-x^2)}\sqrt{[1+\{x/\sqrt{(a^2-x^2)}\}^2]}\,dx = 2\pi a(p_2-p_1).$$

10.7. Tangent, normal, subtangent, and subnormal

If the tangent at a point P, of a curve, whose coordinates are (X, Y) meets the x-axis at the point T then the distance between the point T and the point $(X, 0)$ is called the length of the *subtangent* of the point P.

The line through P perpendicular to the tangent at P is called the *normal* at P and if the normal meets the x-axis at N the distance between the points N and $(X, 0)$ is called the length of the *subnormal* of P.

Thus if (X, Y) is the point P on the curve $y = f(x)$, since the equation of the tangent at P is $y-Y = (x-X)f'(X)$, therefore the equation of the normal is

$$(y-Y)f'(X)+(x-X) = 0$$

and the coordinates of T and N are

$$(X-Y/f'(X), 0) \quad \text{and} \quad (X+Yf'(X), 0).$$

It follows that the length of the subtangent of P is

$$|Y/f'(X)|$$

and the length of the subnormal is

$$|Yf'(X)|.$$

EXAMPLE. The length of the subtangent of the point (x, y) of the parabola $y^2 = 2ax$ is

$$\left|y\Big/\frac{dy}{dx}\right| = \left|y\frac{dx}{dy}\right| = \left|\frac{y^2}{a}\right| = 2|x|$$

and the length of the subnormal is

$$\left|y\frac{dy}{dx}\right| = \left|y\Big/\frac{dx}{dy}\right| = |a|$$

Thus the subnormal of the parabola is of constant length; we shall see in the next chapter that the parabola is the only curve with constant subnormal.

10.71. If $x = X$, $y = Y$ when $t = T$ then the equation of the tangent to the curve $x = x(t)$, $y = y(t)$ at the point $t = T$ is

$$\frac{x-X}{X'} = \frac{y-Y}{Y'}, \quad \text{where } X' = x'(T),\ Y' = y'(T),$$

and so the equation of the normal takes the form

$$(x-X)X'+(y-Y)Y' = 0.$$

10.72. If ψ is the inclination of the tangent to the curve $y = f(x)$ at the point (x, y) then, by 10.21,

$$\tan\psi = f'(x).$$

But if $s(x)$ is the length of the arc of the curve from a to x then

$$s(x) = \int_a^x +\sqrt{[1+\{f'(x)\}^2]}\,dx$$

and therefore

$$\frac{ds}{dx} = +\sqrt{[1+\{f'(x)\}^2]} = +\sqrt{(1+\tan^2\psi)} = \sec\psi$$

(since $-\tfrac{1}{2}\pi < \psi < \tfrac{1}{2}\pi$),

so that $$\frac{dx}{ds} = \cos\psi,$$

and $$\frac{dy}{ds} = \frac{dy}{dx}\frac{dx}{ds} = \tan\psi\cos\psi = \sin\psi,$$

i.e. $$\frac{dy}{ds} = \sin\psi.$$

10.73. If p is the length of the perpendicular from the origin to the tangent to the curve $y = f(x)$ at the point (X, Y) then

$$p = |Y\cos\psi - X\sin\psi|,$$

where ψ is the inclination of the tangent.

For the equation of the tangent is

$$(y-Y)-(x-X)\tan\psi = 0,$$

i.e. $$y - x\tan\psi - (Y - X\tan\psi) = 0,$$

and therefore

$$p = \left|\frac{Y - X\tan\psi}{\sqrt{(1+\tan^2\psi)}}\right. \quad : |Y\cos\psi - X\sin\psi|.$$

10.74. If r is the distance from the origin to the point (x, y), on the curve $y = f(x)$, ψ the inclination of the tangent at this point, p the perpendicular distance of the tangent from the origin, and s the length of the arc of the curve then

$$\left|r\frac{dr}{dp}\right. \qquad \left|\overline{d\psi}\right.$$

Let $q = y\cos\psi - x\sin\psi$ so that $p = |q|$.

We observe first that since $p^2 = q^2$ therefore $p\dfrac{dp}{dt} = q\dfrac{dq}{dt}$ and so $p\left|\dfrac{dp}{dt}\right. = |q \left|\dfrac{dq}{dt}\right|$, whence $\left|\dfrac{dp}{dt}\right. \qquad |$, for any t.

Now

$$\frac{dq}{dx} = \frac{dy}{dx}\cos\psi - y\sin\psi\frac{d\psi}{dx} - \sin\psi - x\cos\psi\frac{d\psi}{dx}$$

$$= -(y\sin\psi + x\cos\psi)\frac{d\psi}{dx}, \quad \text{since } \frac{dy}{dx}\cos\psi = \tan\psi\cos\psi = \sin\psi$$

therefore $$\frac{dq}{dr} = \frac{dq}{dx}\frac{dx}{dr} = -(y\sin\psi + x\cos\psi)\frac{d\psi}{dr},$$

whence $$r\frac{dr}{dq} = -r\frac{dr}{d\psi}\Big/(y\sin\psi + x\cos\psi).$$

But $r^2 = x^2 + y^2$ so that

$$r\frac{dr}{d\psi} = x\frac{dx}{d\psi} + y\frac{dy}{d\psi} = \left\{x\frac{dx}{ds} + y\frac{dy}{ds}\right\}\frac{ds}{d\psi}$$

$$= \{x\cos\psi + y\sin\psi\}\frac{ds}{d\psi}$$

$$r\frac{dr}{\ldots} \quad \frac{ds}{\ldots}.$$

10.75. The line joining the point (r, θ) to the origin is called the *radius vector* through the point; its equation is $x \sin\theta = y\cos\theta$. If ϕ^* is the acute angle between the radius vector and the tangent through the point (r, θ), to the curve $r = f(\theta)$ and if ψ is the inclination of the tangent, then since the tangent is parallel to

$$\left|\frac{\cos\theta\cos\psi+\sin\theta\sin\psi}{\sqrt{\{(\sin^2\theta+\cos^2\theta)(\sin^2\psi+\cos^2\psi\ldots}}\right| = |\cos(\theta-\psi)|$$

But $r^2 = x^2+y^2$ and

$$\frac{dr}{\ldots} = \frac{x}{\ldots}\frac{dx}{\ldots}+\frac{y}{\ldots}\frac{dy}{\ldots} = \cos\theta\cos\psi+\sin\theta\sin\psi = \cos(\theta-\psi\ldots$$

Let $\phi = \phi^*$ if dr/ds is positive and $\phi = \pi-\phi^*$ if dr/ds is negative, then in either case $\cos\phi = dr/ds$. ϕ is usually called the angle between the tangent and radius vector; observe however that ϕ is the acute or the obtuse angle between the tangent and radius vector according as dr/ds is positive or negative.

10.8. The curvature of a plane curve

If (X, Y) is a common point of a curve and a circle and if at the point (X, Y) both dy/dx and d^2y/dx^2 have the same value for the curve and the circle then the circle is called the *circle of curvature* of the curve at the point (X, Y); the radius of this circle is called the *radius of curvature* and its centre the *centre of curvature*; the inverse of the radius of curvature is called the *curvature*.

10.81. To find the radius of curvature at the point (X, Y) on the curve $y = f(x)$.

Suppose the equation of the circle of curvature is

$$(x-a)^2+(y-b)^2 = \rho^2. \qquad \text{(i)}$$

then at any point on the circle we have

$$(x-a)+(y-b)\frac{dy}{dx} = 0, \tag{ii}$$

$$1+(y-b)\frac{d^2y}{dx^2}+\left(\frac{dy}{dx}\right)^2 = 0. \tag{iii}$$

The equations (i), (ii), and (iii) must be satisfied by

$$x = X, \qquad y = Y, \qquad \frac{dy}{dx} = f'(X), \qquad \frac{d^2y}{dx^2} = f''(X)$$

therefore $\qquad (X-a)^2+(Y-b)^2 = \rho^2$

and

$$(X-a)+(Y-b)f'(X) = 0, \qquad 1+(Y-b)f''(X)+\{f'(X)\}^2 = 0, \tag{iv}$$

whence

$$\frac{X-a}{f'(X)} = -(Y-b) = \frac{+\sqrt{\{(X-a)^2+(Y-b)^2\}}}{\sqrt{[1+\{f'(X)\}^2]}} = \frac{\rho}{\sqrt{[1+\{f'(X)\}^2]}},$$

the square root having the same sign as $-(Y-b)$, and so the same sign as $f''(X)$, since $-(Y-b)f''(X) = 1+\{f'(X)\}^2 > 0$; hence

$$1+\{f'(X)\}^2 = -(Y-b)f''(X) = \rho f''(X)/\sqrt{[1+\{f'(X)\}^2]},$$

which shows that $\qquad \rho = [1+\{f'(X)\}^2]^{\frac{3}{2}}/f''(X),$

the square root having the same sign as $f''(X)$.

Thus the radius of curvature at the point (x, y) of the curve $y = f'(x)$ is the positive value of

$$[1+\{\mathbf{f}'(\mathbf{x})\}^2]^{\frac{3}{2}}/\mathbf{f}''(\mathbf{x}).$$

10.811. Solving the equations (iv), § 10.81, for a and b, we see that the centre of curvature (a, b) at the point (x, y) on the curve $y = f(x)$ is given by

$$a = x-\{1+f'(x)^2\}f'(x)/f''(x),$$
$$b = y+\{1+f'(x)^2\}/f''(x).$$

10.812. Writing dy/dx for $f'(x)$ and d^2y/dx^2 for $f''(x)$ the formula 10.81 for the radius of curvature becomes

$$\left|\left\{1+\left(\frac{dy}{dx}\right)^2\right\}^{\frac{3}{2}}\Big/\frac{d^2y}{dx^2}\right|$$

This formula is unaltered by interchanging x and y; for

$$\frac{dx}{dy} = 1\Big/\frac{dy}{dx}, \qquad \frac{d^2x}{dy^2} = \frac{d}{dx}\left(1\Big/\frac{dy}{dx}\right)\frac{dx}{dy} = -\frac{d^2y}{dx^2}\left(\frac{dx}{dy}\right)^3$$

$$\left\{1+\left(\frac{dx}{dy}\right)^2\right\}^{\frac{3}{2}} = \left\{1+\left(\frac{dy}{dx}\right)^2\right\}^{\frac{3}{2}}\left(\frac{dx}{dy}\right)^3$$

so that
$$\left\{1+\left(\frac{dx}{dy}\right)^2\right\}^{\frac{3}{2}}\Big/\frac{d^2x}{dy^2} = -\left\{1+\left(\frac{dy}{dx}\right)^2\right\}^{\frac{3}{2}}\Big/\frac{d^2y}{dx^2}$$

and therefore

$$\left|\left\{1+\left(\frac{dx}{dy}\right)^2\right\}^{\frac{3}{2}}\Big/\frac{d^2x}{dy^2}\right| = \left|\left\{1+\left(\frac{dy}{dx}\right)^2\right\}^{\frac{3}{2}}\Big/\frac{d^2y}{dx^2}\right|$$

10.82. The radius of curvature at the point t on the curve $x = x(t)$, $y = y(t)$ is

$$\{x'(t)^2+y'(t)^2\}^{\frac{3}{2}}/\{x'(t)y''(t)-x''(t)y'(t)\}.$$

For if the result of eliminating t between $x = x(t)$, $y = y(t)$ is

$$y = f(x)$$

then $y(t) = f(x(t))$, for all t. Whence

$$y' = f'(x)x'$$

and
$$y'' = f''(x)x'^2+f'(x)x'',$$

so that the radius of curvature is

$$\frac{[1+\{f'(x)\}^2]^{\frac{3}{2}}}{f''(x)} = \frac{\{1+y'^2/x'^2\}^{\frac{3}{2}}}{(x'y''-y'x'')/x'^3} = \frac{\{x'^2+y'^2\}^{\frac{3}{2}}}{x'y''-x''y'},$$

the square root having the same sign as $x'y''-x''y'$.

10.83. To prove that $\rho = |ds/d\psi|$.

Since $f'(x) = \tan\psi$, therefore

$$f''(x) = \sec^2\psi\frac{d\psi}{dx} = \sec^2\psi\frac{ds}{dx}\frac{d\psi}{ds} = \sec^3\psi\frac{d\psi}{ds}$$

and so
$$\rho = [1+\{f'(x)\}^2]^{\frac{3}{2}}/f''(x)$$

$$= [\sec^2\psi]^{\frac{3}{2}}\Big/\sec^3\psi\frac{d\psi}{ds} = \left|\sec^3\psi\Big/\sec^3\psi\frac{d\psi}{ds}\right|$$

whence
$$\rho = \left|\frac{ds}{d\psi}\right|.$$

10.84. The radius of curvature at the point (r, θ) on the curve $r = f(\theta)$ is

$$\{r^2+r'^2\}^{\frac{3}{2}}/\{r^2+2r'^2-rr''\},$$

where $r' = dr/d\theta$, $r'' = d^2r/d\theta^2$.

For at any point of the curve we have

$$x = f(\theta)\cos\theta = r\cos\theta, \qquad y = f(\theta)\sin\theta = r\sin\theta,$$

so that

$$x' = r'\cos\theta - r\sin\theta, \qquad y' = r'\sin\theta + r\cos\theta,$$

$$x'' = r''\cos\theta - 2r'\sin\theta - r\cos\theta, \qquad y'' = r''\sin\theta + 2r'\cos\theta - r\sin\theta,$$

whence the result stated follows from 10.82 after a simple calculation.

10.85. From 10.83 and 10.74 we have $\rho = |r(dr/dp)|$.

10.86. To prove that

$$\frac{1}{\rho^2} = \left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2.$$

We have

$$\frac{dx}{ds} = \cos\psi \quad \text{and so} \quad \frac{d^2x}{ds^2} = -\sin\psi\frac{d\psi}{ds}.$$

Similarly
$$\frac{d^2y}{ds^2} = \cos\psi\frac{d\psi}{ds}$$

and therefore

$$\left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2 = (\sin^2\psi+\cos^2\psi)\left(\frac{d\psi}{ds}\right)^2 = \frac{1}{\rho^2} \quad \text{by 10.83.}$$

10.9. Definitions. If $Y > y$ then the point (X, Y) is said to be *above* the point (x, y), and the latter is said to be *below* the former.

10.901. A point (X, Y) is said to be *above* a curve $y = f(x)$ if $Y > f(X)$, and *below* if $Y < f(X)$. Two points are said to be *on the same side of* a curve if they are both above or both below the curve.

Observe that if (X, Y) is above $y = f(x)$, then provided $f(x)$ is continuous, (X, Y) is above all points (x, y) of the curve for which x is sufficiently close to X; for when $|X-x|$ is small enough $|f(x)-f(X)| < Y-f(X)$ and so $f(x) < Y$.

10.902. A curve $y = f(x)$ is said to be *above* a curve $y = g(x)$, in (a, b), if $f(x) > g(x)$ in (a, b), and then $y = g(x)$ is said to be *below* $y = f(x)$.

If two curves c_1 and c_2 are both above, or both below, a curve c they are said to be on the same side of c.

If $y = f(x)$ is above $y = g(x)$, and $f(x)$ and $g(x)$ are continuous then any point (x, y) on the first curve is above any point (X, Y) of the second, provided x and X are sufficiently close. For by hypothesis $f(x) > g(x)$ and so the point $(x, f(x))$ is above the curve $y = g(x)$, and therefore, by 10.901, above any point (X, Y) on $y = g(x)$ for which $|X-x|$ is sufficiently small.

10.91. *If $f''(a) > 0$ the centre of curvature of the curve $y = f(x)$ at the point $x = a$ is above the tangent to the curve at this point, and if $f''(a) < 0$ the centre of curvature is below the tangent.*

The centre of curvature at the point $(a, f(a))$ has the coordinates

$$x_a = a-\{1+f'(a)^2\}f'(a)/f''(a), \qquad y_a = f(a)+\{1+f'(a)^2\}/f''(a).$$

The tangent at $x = a$ is $y = f(a)+(x-a)f'(a)$ and so when $x = x_a$, on the tangent,

$$y = y_t = f(a)-\{1+f'(a)^2\}f'(a)^2/f''(a).$$

Hence $$y_a-y_t = \{1+f'(a)^2\}^2/f''(a),$$

and therefore if $f''(a) > 0$, $y_a > y_t$ and the centre of curvature is above the tangent, and if $f''(a) < 0$, $y_a < y_t$ and the centre of curvature is below the tangent.

10.92. *If $f''(a) > 0$, then in the neighbourhood of the point a the curve $y = f(x)$ lies above the tangent to the curve at this point.*

First we prove that the function $\{f(x)-f(a)\}/(x-a)$ is *increasing* in the neighbourhood of a.

The derivative of $\{f(x)-f(a)\}/(x-a)$ is

$$\{(x-a)f'(x)-(f(x)-f(a))\}/(x-a)^2, \quad x \neq a,$$

which has the same sign as $\phi(x) = (x-a)f'(x)-\{f(x)-f(a)\}$. But $\phi'(x) = (x-a)f''(x)$ and for values of x near a, $''(x)$ being continuous has the same sign as $f''(a)$, and therefore $\phi(x)$ decreases for $x < a$ and increases for $x > a$. Since $\phi(a) = 0$, it follows that $\phi(x) > 0$, $x \neq a$. Hence $\{f(x)-f(a)\}/(x-a)$ is increasing both for

$x < a$ and $x > a$. Furthermore $\{f(x_n)-f(a)\}/(x_n-a) \to f'(a)$ as $x_n \to a$ and so

$$\{f(x^*)-f(a)\}/(x^*-a) < f'(a) < \{f(x)-f(a)\}/(x-a), \quad x^* < a < x,$$

whence $$f(x) > f(a)+(x-a)f'(a)$$

whether x be greater or less than a. Thus the point $(x, f(x))$ on the curve lies above the point $(x, f(a)+(x-a)f'(a))$ on the tangent to the curve at the point $(a, f(a))$, for all values of x near a.

Similarly, if $f''(a) < 0$ the curve lies below the tangent.

10.921. *Theorems* 10.91 *and* 10.92 *together show that at any point of a curve the centre of curvature is on the same side of the tangent as the curve.*

10.93. Theorem 10.92 may also be expressed in the following way:

If $f''(x) > 0$ *throughout an interval* (a, b) *then in this interval the curve* $y = f(x)$ *lies below the chord joining the points* $(a, f(a))$, $(b, f(b))$ *on the curve.*

For $\{f(x)-f(a)\}/(x-a)$ is increasing in $[a, b)$ and so if x lies in $[a, b]$ then

$$\{f(x)-f(a)\}/(x-a) < \{f(b)-f(a)\}/(b-a),$$

and so $$f(x) < f(a)+\{f(b)-f(a)\}(x-a)/(b-a),$$

i.e. $$f(x) < \{(b-x)f(a)+(x-a)f(b)\}/[(b-x)+(x-a)],$$

which shows that the point $(x, f(x))$ on the curve lies below the point $x, \{(b-x)f(a)+(x-a)f(b)\}/(b-a)$ on the line joining $(a, f(a))$ to $(b, f(b))$.

Similarly, if $f''(x) < 0$ in (a, b) then in this interval the arc of the curve $y = f(x)$ lies above the chord joining its end-points.

10.931. If $f''(a) = 0$ and $f''(x)$ changes sign as x passes through the value a, then the curve $y = f(x)$ crosses the tangent at $x = a$.

For the difference between the y-coordinates of the point $x, f(x)$ on the curve and the point $x, f(a)+(x-a)f'(a)$ on the tangent is

$$\phi(x) = f(x)-f(a)-(x-a)f'(a).$$

Since $\phi'(x) = f'(x)-f'(a)$ and $\phi''(x) = f''(x)$, therefore all three $\phi(x)$, $\phi'(x)$, and $\phi''(x)$ vanish at $x = a$, and $\phi''(x)$ changes sign as x passes through a. Hence $\phi'(x)$ either first increases to $\phi'(a) = 0$

and then decreases, or vice versa, so that $\phi'(x)$ is of constant sign near $x = a$, and therefore $\phi(x)$ either steadily increases or steadily decreases through the value $\phi(a) = 0$, so that $\phi(x)$ changes sign as x passes through the value a. Thus the curve crosses its tangent at $x = a$.

A point of a curve where $f''(x)$ vanishes and changes sign is called a *point of inflexion*, and the tangent at such a point is called an *inflexion tangent*. An inflexion tangent has a closer relationship to the curve than an ordinary tangent, since not only dy/dx but also d^2y/dx^2 has the same value, at the point of contact, on the curve and on the tangent.

If $f''(a) = 0$ but $f''(x)$ does not change sign as x passes through a, then the curve does *not* cross the tangent at $x = a$.

For when $\phi''(x)$ vanishes at $x = a$ without changing sign, then $\phi'(x)$ changes sign as x passes through a, and therefore $\phi(x)$ itself does *not* change sign.

10.94. If the centres of curvature of a curve C compose a curve C' then C' is called the *evolute* of the curve C, and C is called an *involute* of C'.

Thus, by 10.811, the parametric equations of the evolute of the curve $y = f(x)$ are

$$X = x - \{1+f'(x)^2\}f'(x)/f''(x), \qquad Y = y + \{1+f'(x)^2\}/f''(x),$$

and so the evolute of $x = x(t)$, $y = y(t)$ has the equations

$$X = x - (x'^2+y'^2)y'/(x'y''-x''y'),$$
$$Y = y + (x'^2+y'^2)x'/(x'y''-x''y'),$$

where x' stands for dx/dt, etc., since $dy/dx = y'/x'$ and

$$\frac{d^2y}{dx^2} = (x'y''-x''y')/x'^3$$

10.95. The parametric equations of the involutes of the curve $x = x(t)$, $y = y(t)$ are, with an arbitrary constant a,

$$\xi = x - \left\{\int_a^t (x'^2+y'^2)^{\frac{1}{2}}\,dt\right\} x'/(x'^2+y'^2)^{\frac{1}{2}},$$

$$\eta = y - \left\{\int_a^t (x'^2+y'^2)^{\frac{1}{2}}\,dt\right\} y'/(x'^2+y'^2)^{\frac{1}{2}}.$$

Proof. $\xi' = x' - x'(x'^2+y'^2)^{\frac{1}{2}}/(x'^2+y'^2)^{\frac{1}{2}} -$

$$-x''\left\{\int (x'^2+y'^2)^{\frac{1}{2}}\,dt\right\}\Big/(x'^2+y'^2)^{\frac{1}{2}} +$$

$$+x'(x'x''+y'y'')\left\{\int (x'^2+y'^2)^{\frac{1}{2}}\,dt\right\}\Big/(x'^2+y'^2)^{\frac{3}{2}}$$

$$= \left\{\int (x'^2+y'^2)^{\frac{1}{2}}\,dt\right\} y'(x'y''-x''y')/(x'^2+y'^2)^{\frac{3}{2}}$$

and $\eta' = -\left\{\int (x'^2+y'^2)^{\frac{1}{2}}\,dt\right\} x'(x'y''-x''y')/(x'^2+y'^2)^{\frac{3}{2}},$

and so $\xi'x'+\eta'y' = 0$, whence

$$\xi'' = -(y'/x')\eta''+(x''y'-x'y'')\eta'/x'^2$$

and therefore

$$\xi''\eta'-\xi'\eta'' = (x''y'-x'y'')\eta'^2/x'^2.$$

Moreover

$$\xi'^2+\eta'^2 = \{(x'^2+y'^2)/x'^2\}\eta'^2.$$

Hence the parametric equations of the evolute of the curve (ξ, η) are

$$X = \xi-(x'^2+y'^2)\eta'/(x'y''-x''y') = \xi+\left\{\int (x'^2+y'^2)\,dt\right\}x'/(x'^2+y'^2)^{\frac{1}{2}}$$

$$= \xi+(x-\xi) = x$$

and $Y = \eta+(x'^2+y'^2)\xi'/(x'y''-x''y') = \eta+(y-\eta) = y,$

i.e. the evolute is $X = x(t)$, $Y = y(t)$, which completes the proof.

10.96. If E is the centre of curvature at the point P of a curve C then the normal at P, to the curve C, is the tangent to the evolute of C, at E.

Let X, Y be the centre of curvature at the point $x(t)$, $y(t)$ on the curve $x = x(t)$, $y = y(t)$; we prove first that

(i) $(X-x)x'+(Y-y)y' = 0,$

(ii) $(X-x)x''+(Y-y)y'' = x'^2+y'^2.$

By § 10.94,

$$(X-x)x'+(Y-y)y' = (x'^2+y'^2)\{x'y'-x'y'\}/(x'y''-x''y') = 0,$$

and

$$(X-x)x''+(Y-y)y'' = (x'^2+y'^2)\{x'y''-x''y'\}/(x'y''-x''y') = x'^2+y'^2.$$

Differentiating (i) it follow that

$$X'x'+Y'y' = 0.$$

The normal at (x, y) is $(\xi-x)x'+(\eta-y)y' = 0$, and since

$$(X-x)x'+(Y-y)y' = 0,$$

the equation of the normal may be written in the form

$$(\xi-X)x'+(\eta-Y)y' = 0,$$

and since $X'x'+Y'y' = 0$ this is the same as

$$(\xi-X)/X' = (\eta-Y)/Y',$$

which is the tangent to the evolute at (X, Y).

10.97. If E_0, E_1 are the centres of curvature at the points P_0, P_1 on a curve C, and R_0, R_1 the radii of curvature at these points, then the length of the arc of the evolute of C, from E_0 to E_1, is $|R_1-R_0|$, provided R' is of constant sign between P_0 and P_1.

Since the tangent to the evolute at (X, Y) passes through (x, y) we have

$$X'/(X-x) = Y'/(Y-y). \qquad \text{(i)}$$

Differentiating the identity

$$(X-x)^2+(Y-y)^2 = R^2 \qquad \text{(ii)}$$

we have $$(X-x)X'+(Y-y)Y' = RR', \qquad \text{(iii)}$$

whence from (i), (ii), and (iii)

$$R^2X' = RR'(X-x)$$

or $$X' = (X-x)R'/R, \quad \text{provided } R \neq 0,$$

and similarly $$Y' = (Y-y)R'/R,$$

whence $$\sigma' = +\sqrt{(X'^2+Y'^2)} = |R'|,$$

where σ is the arc-length of the evolute. Hence if R' is positive between P_0, P_1 we have

$$\sigma = \int^{t_1} R'\,dt = R_1-R_0,$$

and if R' is negative,

$$= -\int^{t_1} R'\,dt = R_0-R_1,$$

which completes the proof.

XI

DIFFERENTIAL EQUATIONS

LINEAR EQUATIONS WITH CONSTANT COEFFICIENTS. SIMULTANEOUS LINEAR EQUATIONS

11. A differential equation is a functional relation between a variable x, some function y, and one or more derivatives of y with respect to x. For instance

$$xy\frac{d^2y}{dx^2}-(x^2+1)\frac{dy}{dx}=0 \quad \text{and} \quad \left(\frac{dy}{dx}\right)^2+y\frac{d^2y}{dx^2}=0$$

are differential equations. *Integration* of a differential equation consists in the determination of a relation between x and y which does not contain any of the derivatives and which satisfies the differential equation; thus a solution of the equation $y\dfrac{dy}{dx}-x=0$ is $y^2-x^2=a$, where a is any constant, for if y^2-x^2 is constant then $\dfrac{d}{dx}(y^2-x^2)=0$, i.e. $y\dfrac{dy}{dx}-x=0$, and so the differential equation is satisfied.

11.01. Differential equations arise in the attempt to solve many diverse problems; commonest of these are problems in Dynamics, but many problems in the Calculus also give rise to differential equations. For example, if we seek to find for what curves the angle between the tangent and chord has the constant value α, we must solve the equation $r\dfrac{d\theta}{dr}=\tan\alpha$; writing the equation in the form $\dfrac{1}{r}\dfrac{dr}{d\theta}-\cot\alpha=0$ the left-hand side is seen to be the derivative of

$$\int\left\{\frac{1}{r}\frac{dr}{d\theta}-\cot\alpha\right\}d\theta=\int\frac{1}{r}\frac{dr}{d\theta}\,d\theta-\theta\cot\alpha=\log r-\theta\cot\alpha,$$

and so a solution is

$$\log r-\theta\cot\alpha=\text{constant}=\log c \quad \text{(say)}$$

and so $$r=ce^{\theta\cot\alpha}$$

which is the equation of the *equiangular spiral.*

11.1. The first type of differential equation we shall consider takes the form

$$M\frac{dy}{dx} = N,$$

where M is a function of y alone and N is a function of x alone.

Since the derivative of $\int\left\{M\frac{dy}{dx}-N\right\}dx$ is $M\frac{dy}{dx}-N$ it follows that the solution of the equation is

$$\int\left\{M\frac{dy}{dx}-N\right\}dx = \text{constant} = a \quad \text{(say)},$$

i.e.
$$\int M\frac{dy}{dx}\,dx = \int N\,dx+a,$$

and so
$$\int M\,dy = \int N\,dx+a.$$

EXAMPLE. The solution of $\dfrac{2\log y}{y}\dfrac{dy}{dx} = \cos x$ is

$$\int\frac{2\log y}{y}\,dy = \int\cos x\,dx+a, \quad \text{i.e.} \quad (\log y)^2 = \sin x+a.$$

11.2. *The equation $dy/dx+Py = Q$, where P, Q are functions of x alone.*

If we can find R, a function of x alone, such that $dR/dx = PR$, then

$$R\frac{dy}{dx}+PRy = R\frac{dy}{dx}+\frac{dR}{dx}y = \frac{d}{dx}(Ry)$$

and the differential equation becomes

$$\frac{d}{dx}(Ry) = RQ,$$

of which the solution is

$$Ry = \int RQ\,dx+a.$$

It remains to show that we can find an R satisfying $dR/dx = PR$; writing this equation in the form

$$\frac{1}{R}\frac{dR}{dx} = P,$$

we see
$$\log R = \int P\,dx,$$

and therefore
$$R = e^{\int P\,dx}.$$

Since the solution of the differential equation is effected by multiplying the equation by this R, the function R is called an *integrating factor* of the differential equation.

EXAMPLE. To solve $x\log x\dfrac{dy}{dx}+y = \log x$.

Write the equation in the form

$$\frac{dy}{dx}+\frac{1}{x\log x}y=\frac{1}{x}$$

The integrating factor is

$$e^{\int 1/(x\log x)\,dx}=e^{\log(\log x)}=\log x$$

and the solution is

$$y\log x = \int \frac{\log x}{x}\,dx+a,$$

i.e. $$y\log x=\tfrac{1}{2}(\log x)^2+a.$$

11.3. Linear equations with constant coefficients

$$a_n\frac{d^ny}{dx^n}+a_{n-1}\frac{d^{n-1}y}{dx^{n-1}}+\dots+a_0y = R,$$

each a_r constant, R a function of x.

Writing D for d/dx the equation takes the form

$$a_nD^ny+a_{n-1}D^{n-1}y+\dots+a_0y=R,$$

which we shall abbreviate to

$$(a_nD^n+a_{n-1}D^{n-1}+\dots+a_0)y=R,\quad\text{or}\quad\Big(\sum_{r=0}^{n}a_rD^r\Big)y=R.$$

If $L(t)$ denotes the polynomial $a_nt^n+a_{n-1}t^{n-1}+\dots+a_1t+a_0$ the equation may be further abbreviated to $L(D)y=R$. When $L(t)$ is of the nth degree we say that the differential equation $L(D)y=R$ is of the nth order.

Before discussing the solution of this equation we shall prove a number of simple results on which the method of solution is based.

11.31. If y_1 and y_2 are any two functions of x then

$$L(D)(y_1+y_2)=L(D)y_1+L(D)y_2.$$

For $D(y_1+y_2)=Dy_1+Dy_2$ and if $D^k(y_1+y_2)=D^ky_1+D^ky_2$ then

$$\begin{aligned}D^{k+1}(y_1+y_2)&=D\{D^k(y_1+y_2)\}=D\{D^ky_1+D^ky_2\}\\&=D.D^ky_1+D.D^ky_2=D^{k+1}y_1+D^{k+1}y_2\end{aligned}$$

and so $$D^r(y_1+y_2) = D^r y_1 + D^r y_2 \quad \text{for any } r.$$

Hence $$\sum_0^n a_r D^r(y_1+y_2) = \sum_0^n a_r D^r y_1 + \sum_0^n a_r D^r y_2,$$

i.e. $$L(D)(y_1+y_2) = L(D)y_1 + L(D)y_2.$$

11.32. If a is constant and y a function of x then $L(D)ay = aL(D)y$ (proof as in 11.31).

11.33. $D^m(D^n y) = D^n(D^m y)$, for any m and n, since each side of the equation is equal to $D^{m+n}y$.

11.331. If $f(t)$ and $g(t)$ denote the polynomials
$$a_m t^m + a_{m-1}t^{m-1} + \ldots + a_0 \quad \text{and} \quad b_n t^n + b_{n-1}t^{n-1} + \ldots + b_0$$
then $$f(D)\{g(D)y\} = g(D)\{f(D)y\} = (f(D)g(D))y.$$

For by 11.32 and 11.33
$$a_r D^r(b_s D^s y) = b_s D^s(a_r D^r y) = (a_r D^r)(b_s D^s)y$$
and so
$$\sum_{r=0}^m a_r D^r(b_s D^s y) = \sum_{r=0}^m b_s D^s(a_r D^r y) = \sum_{r=0}^m (a_r D^r)(b_s D^s)y,$$
i.e. $f(D)(b_s D^s y) = b_s D^s(f(D)y) = (f(D)b_s D^s)y$, using 11.31,

and adding these equations for $s = 0, 1, 2, \ldots, n$ we find
$$f(D)\{g(D)y\} = g(D)\{f(D)y\} = \{f(D)g(D)\}y.$$

11.332. If $L(t)$ denotes the polynomial
$$a_n t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0$$
and $L'(t), L''(t), \ldots, L^r(t), \ldots$ its successive derivatives, and if u, v are any two functions of x, then
$$L(D)uv = uL(D)v + Du\,L'(D)v + \\ + \frac{D^2u}{2!}L''(D)v + \ldots + \frac{D^r u}{r!}L^r(D)v + \ldots + \frac{D^n u}{n!}L^n(D)v.$$

Proof. The Leibnitz formula for the rth derivative of uv may be written in the form
$$D^r uv = uD^r v + Du\,rD^{r-1}v + \frac{D^2u}{2!}r(r-1)D^{r-2}v + \\ + \frac{D^3u}{3!}r(r-1)(r-2)D^{r-3}v + \ldots + \frac{D^r u}{r!}r!\,v$$

and therefore

$$L(D)uv = \sum_0^n a_r D^r uv$$

$$= \sum_0^n \Big\{ua_r D^r v + Du\, ra_r D^{r-1}v + \frac{D^2u}{2!}r(r-1)a_r D^{r-2}v + \dots +$$

$$+\frac{D^k u}{k!}r(r-1)\dots(r-k+1)a_r D^{r-k}v + \dots + \frac{D^r u}{r!}r!\, a_r v\Big\}$$

$$= uL(D)v + Du\, L'(D)v + \frac{D^2u}{2!}L''(D)v + \dots +$$

$$+\frac{D^k u}{k!}L^k(D)v + \dots + \frac{D^n u}{n!}L^n(D)v.$$

11.4. $\qquad \mathbf{L(D)e^{px} = e^{px}L(p).}$

For $D^n e^{px} = p^n e^{px}$, and so $\sum a_r D^r e^{px} = \sum a_r p^r e^{px}$, i.e.

$$L(D)e^{px} = L(p).e^{px}.$$

11.41. $\qquad \mathbf{L(D)e^{px}y = e^{px}L(D+p)y.}$

For, by formula 11.332,

$$L(D)e^{px}y = e^{px}\Big\{L(D) + pL'(D) + \frac{p^2}{2!}L''(D) + \dots + \frac{p^n}{n!}L^n(D)\Big\}y$$

$$= e^{px}L(p+D)y, \quad \text{by Example 7.92.}$$

11.411. $\qquad e^{px}L(D)y = L(D-p)e^{px}y,$

for $\qquad L(D-p)e^{px}y = e^{px}L(D+p-p)y = e^{px}L(D)y.$

11.42. $\qquad \mathbf{L(D^2)\sin px = L(-p^2)\sin px}$

and $\qquad \mathbf{L(D^2)\cos px = L(-p^2)\cos px.}$

For $D^2\sin px = -p^2\sin px$ and so if $(D^2)^n\sin px = (-p^2)^n\sin px$ then

$$(D^2)^{n+1}\sin px = D^2\{(D^2)^n\sin px\} = D^2\{(-p^2)^n\sin px\}$$

$$= (-p^2)^{n+1}\sin px,$$

so that $(D^2)^r\sin px = (-p^2)^r\sin px$ for all values of r.

Hence

$$L(D^2)\sin px = \sum a_r(D^2)^r\sin px = \sum a_r(-p^2)^r\sin px = L(-p^2)\sin px.$$

Similarly $\qquad L(D^2)\cos px = L(-p^2)\cos px.$

EXAMPLES. $(D^2+2D+2)e^{3x} = (3^2+2.3+2)e^{3x} = 17e^{3x}$.

$$(D^2+2D+2)e^{3x}x^2 = e^{3x}\{(D+3)^2+2(D+3)+2\}x^2$$
$$= e^{3x}\{D^2+8D+17\}x^2 = e^{3x}\{2+16x+17x^2\}.$$

$$(D^4+D^2+1)(A\sin 2x+B\cos 3x)$$
$$= A\{(-2^2)^2-2^2+1\}\sin 2x+B\{(-3^2)^2-3^2+1\}\cos 3x$$
$$= 13A\sin 2x+73B\cos 3x.$$

11.43. $$(D^2+a^2)^n x^n\cos ax = (2a)^n n!\cos(ax+\tfrac{1}{2}n\pi).$$

11.431. $$(D^2+a^2)^n x^n\sin ax = (2a)^n n!\sin(ax+\tfrac{1}{2}n\pi).$$

Proof. Since

$$(D^2+a^2)\cos ax = (-a^2+a^2)\cos ax = 0,$$

by 11.332 we have

$$(D^2+a^2)x\cos ax = x(D^2+a^2)\cos ax+2D\cos ax = -2a\sin ax,$$

i.e. $$(D^2+a^2)x\cos ax = 2a\cos(ax+\tfrac{1}{2}\pi).$$

Similarly $$(D^2+a^2)x\sin ax = 2a\sin(ax+\tfrac{1}{2}\pi).$$

Thus 11.43 and 11.431 are true for $n = 1$. Suppose that they are true for $n = k$. Then

$$(D^2+a^2)^{k+1}x^k\cos ax = (D^2+a^2)(2a)^k k!\cos(ax+\tfrac{1}{2}k\pi) = 0,$$

and so, by 11.332,

$$(D^2+a^2)^{k+1}x^{k+1}\cos ax = (D^2+a^2)^{k+1}x\,x^k\cos ax$$
$$= x(D^2+a^2)^{k+1}x^k\cos ax+2(k+1)D(D^2+a^2)^k x^k\cos ax$$
$$= -2a(k+1)(2a)^k k!\sin(ax+\tfrac{1}{2}k\pi),$$

whence

$$(D^2+a^2)^{k+1}x^{k+1}\cos ax = (2a)^{k+1}(k+1)!\cos\{ax+\tfrac{1}{2}(k+1)\pi\}$$
$$\text{since}\quad -\sin(ax+\tfrac{1}{2}k\pi) = \cos\{ax+\tfrac{1}{2}(k+1)\pi\},$$

and similarly

$$(D^2+a^2)^{k+1}x^{k+1}\sin ax = (2a)^{k+1}(k+1)!\sin\{ax+(\tfrac{1}{2}k+1)\pi\},$$

so that 11.43 and 11.431 are true for $n = k+1$ and so are true for any value of n.

11.44. Since $(D^2+a^2)\cos(ax+\tfrac{1}{2}n\pi) = 0$ it follows from 11.43 that, if $p > n$,

$$(D^2+a^2)^p x^n\cos ax = 0,$$

and similarl $$(D^2+a^2)^p x^n\sin ax = 0.$$

11.45. Theorems 11.43 and 11.431 may be written

$$(D^2+a^2)^n x^n \cos(ax+\tfrac{1}{2}n\pi) = (2a)^n n! \cos ax,$$

$$(D^2+a^2)^n x^n \sin(ax+\tfrac{1}{2}n\pi) = (2a)^n n! \sin ax,$$

for

$$(D^2+a^2)^n x^n \cos(ax+\tfrac{1}{2}n\pi)$$

$$= \cos\tfrac{1}{2}n\pi(D^2+a^2)^n x^n \cos ax - \sin\tfrac{1}{2}n\pi(D^2+a^2)^n x^n \sin ax$$

$$= (2a)^n n!\{\cos(ax+\tfrac{1}{2}n\pi)\cos\tfrac{1}{2}n\pi - \sin(ax+\tfrac{1}{2}n\pi)\sin\tfrac{1}{2}n\pi\}$$

$$= (2a)^n n! \cos ax, \quad \text{etc.}$$

Instead of $\cos(ax+\tfrac{1}{2}n\pi)$ we may of course write $\cos(ax-\tfrac{1}{2}n\pi)$, and $\sin(ax-\tfrac{1}{2}n\pi)$ for $\sin(ax+\tfrac{1}{2}n\pi)$.

11.46.* $u_0, v_0, w_0, \ldots$ are n differentiable functions of x, and t_0 is another; $t_r, u_r, v_r, \ldots$ are the rth derivatives of $t_0, u_0, v_0, \ldots$ respectively.

Then if the determinant

$$\Delta = \begin{vmatrix} t_0 & u_0 & v_0 & w_0 & . & . \\ t_1 & u_1 & v_1 & w_1 & . & . \\ t_2 & u_2 & v_2 & w_2 & . & . \\ . & . & . & . & . & . \\ t_n & u_n & v_n & w_n & . & . \end{vmatrix}$$

is zero for all values of x

but the determinant

$$\delta = \begin{vmatrix} u_0 & v_0 & w_0 & . & . & . \\ u_1 & v_1 & w_1 & . & . & . \\ u_2 & v_2 & w_2 & . & . & . \\ . & . & . & . & . & . \\ u_{n-1} & v_{n-1} & w_{n-1} & . & . & . \end{vmatrix}$$

is different from zero for all values of x,

we shall prove that there are constants

$$A, \quad B, \quad C, \quad \ldots$$

such that $$t_0 = Au_0+Bv_0+Cw_0+\ldots.$$

The determinant Δ is called the Wronskian of the $n+1$ functions $t_0, u_0, v_0, \ldots$; accordingly the determinant δ is the Wronskian of the n functions $u_0, v_0, w_0, \ldots$.

Since δ is not zero the equations

$$\left.\begin{aligned} t_0 &= Ku_0+Lv_0+Mw_0+\ldots, \\ t_1 &= Ku_1+Lv_1+Mw_1+\ldots, \\ &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\ t_{n-1} &= Ku_{n-1}+Lv_{n-1}+Mw_{n-1}+\ldots \end{aligned}\right\} \tag{α}$$

determine unique functions K, L, M,....

$$Ku_n+Lv_n+Mw_n+\ldots = t_n+\sigma. \tag{α'}$$

Eliminating K, L, M,... from the equations (α) and (α'), we have

$$\begin{vmatrix} u_0 & \ldots & w_0 & \\ u_1 & & w_1 & \\ u_2 & & w_2 & \\ \cdot & \cdot & \cdot & \\ t_n+\sigma & u_n & & \end{vmatrix} = 0.$$

i.e. $\Delta+\sigma\delta = 0$, whence $\sigma = 0$, since $\Delta = 0$ and $\delta \neq 0$. Therefore

$$t_n = Ku_n+Lv_n+Mw_n+\ldots. \tag{β}$$

Differentiating the equations (α) and taking into account the equations (α) and (β), we have

$$\begin{aligned} 0 &= K'u_0+L'v_0+M'w_0+\ldots, \\ 0 &= K'u_1+L'v_1+M'w_1+\ldots, \\ 0 &= K'u_2+L'v_2+M'w_2+\ldots, \\ & \\ 0 &= K'u_{n-1}+L'v_{n-1}+M'w_{n-1}+\ldots, \end{aligned}$$

where K', L', M',... are the derivatives of K, L, M,... respectively. Multiplying these equations in turn by the cofactors of u_0, u_1, u_2,... in δ, and adding, we find $K'\delta = 0$, and so $K' = 0$ i.e. K is a constant; similarly L, M,... are constants, which completes the proof.

11.5.* If $y = u_0$, $y = v_0$, $y = w_0$,... are n solutions of the differential equation of the nth order $L(D)y = 0$, then

$$y = Au_0+Bv_0+Cw_0+\ldots$$

is the *general* solution of the equation, i.e. *every* solution of the equation can be obtained from $Au_0+Bv_0+Cw_0+\ldots$ by a suitabl

choice of the constants A, B, C,... provided that the determinant

$$\delta = \begin{vmatrix} u_0 & v_0 & w_0 & . & . \\ u_1 & v_1 & w_1 & . & . \\ u_2 & v_2 & w_2 & . & . \\ . & . & . & . & . \\ u_{n-1} & v_{n-1} & w_{n-1} & . & . \end{vmatrix}$$

is different from zero for all values of x.

A set of n solutions, of an nth order equation, with non-zero Wronskian is called a fundamental set of solutions

Let the differential equation be

$$(a_n D^n + a_{n-1} D^{n-1} + ... + a_0)y = 0$$

and let t_0 be *any* solution. Then

$$\begin{aligned} a_n t_n + a_{n-1} t_{n-1} + ... + a_0 t_0 &= 0, \\ a_n u_n + a_{n-1} u_{n-1} + ... + a_0 u_0 &= 0, \\ a_n v_n + a_{n-1} v_{n-1} + ... + a_0 v_0 &= 0, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . & \end{aligned}$$

Multiplying these equations in turn by the cofactors of t_n, u_n, v_n,... in Δ, and adding, we find $a_n \Delta = 0$; but $a_n \neq 0$ since the differential equation is of the nth order, and so $\Delta = 0$. Hence, by 11.46, there are constants A, B, C,... such that

$$t_0 = Au_0 + Bv_0 + Cw_0 + \dots.$$

11.51. If $y = y_1$ is a solution of the equation $f(D)y = 0$ and $y = y_2$ is a solution of $g(D)y = 0$ then $y = A_1 y_1 + A_2 y_2$ is a solution of $\{f(D)g(D)\}y = 0$.

For

$$f(D)g(D)\{A_1 y_1 + A_2 y_2\} = g(D)\{f(D)A_1 y_1\} + f(D)\{g(D)A_2 y_2\} = 0.$$

It follows that to solve an equation L(D)y = 0 we need only to solve the equations f(D)y = 0 for each factor f(t) of the polynomial L(t).

11.52. The general solution of $(D-a)y = 0$ is $y = a_0 e^{ax}$ for

$$(D-a)a_0 e^{ax} = (a-a)a_0 e^{ax} = 0.$$

11.53. The general solution of $(D-a)^{k+1}y = 0$ is

$$y = e^{ax}(a_0 + a_1 x + a_2 x^2 + ... + a_k x^k),$$

for $\quad (D-a)^{k+1} e^{ax} x^r = e^{ax} D^{k+1} x^r = 0 \quad \text{if } r \leqslant k.$

11.54. To solve $\{(D-a)^2+b^2\}y = 0$. We have

$$e^{-ax}\{(D-a)^2+b^2\}y = 0$$

and therefore $\{D^2+b^2\}e^{-ax}y = 0$; but

$$(D^2+b^2)\sin bx = (D^2+b^2)\cos bx = 0$$

and so the general solution of the equation is

$$e^{-ax}y = A\cos bx+B\sin bx,$$

i.e.
$$y = e^{ax}(A\cos bx+B\sin bx).$$

11.55. To solve $\{(D-a)^2+b^2\}^{k+1}y = 0$. The equation is equivalent to

$$(D^2+b^2)^{k+1}e^{-ax}y = 0.$$

From 11.44 it follows that

$$(D^2+b^2)^{k+1}x^r(A_r\cos bx+B_r\sin bx) = 0, \quad \text{provided } r \leqslant k,$$

and so the general solution is

$$e^{-ax}y = \sum_{r=0}^{n}(A_r\cos bx+B_r\sin bx)x^r,$$

i.e.

$$y = e^{ax}\sum_{r=0}^{n}x^r(A_r\cos bx+B_r\sin bx)$$
$$= e^{ax}\cos bx(A_0+A_1x+\ldots+A_kx^k)+e^{ax}\sin bx(B_0+B_1x+\ldots+B_kx^k).$$

EXAMPLES. (i) $(D^2+D-2)y = 0$. The factors of D^2+D-2 are $D-1$ and $D+2$; the general solutions of $(D-1)y = 0$ and $(D+2)y = 0$ are $y = Ae^x$ and $y = Be^{-2x}$ and so the general solution of $(D^2+D-2)y = 0$ is

$$y = Ae^x+Be^{-2x}.$$

(ii) $(D^3+2D^2+2D+1)y = 0$. The factors of D^3+2D^2+2D+1 are $D+1$ and D^2+D+1.

The solution of $(D+1)y = 0$ is $y = Ae^{-x}$. To solve

$$(D^2+D+1)y = 0$$

we write the equation in the form $\{(D+\frac{1}{2})^2+\frac{3}{4}\}y = 0$, whence

$$(D^2+\tfrac{3}{4})e^{\frac{1}{2}x}y = 0$$

and so the solution is $y = e^{-\frac{1}{2}x}(P\cos\frac{1}{2}\sqrt{3}\,x+Q\sin\frac{1}{2}\sqrt{3}\,x)$; therefore the solution of $(D^3+2D^2+2D+1)y = 0$ is

$$y = Ae^{-x}+e^{-\frac{1}{2}x}(P\cos\tfrac{1}{2}\sqrt{3}\,x+Q\sin\tfrac{1}{2}\sqrt{3}\,x).$$

(iii) $(D+1)^3(D^2-2D+5)^2y = 0$. The solution of $(D+1)^3y = 0$ is $y = e^{-x}(L+Mx+Nx^2)$, and the solution of $(D^2-2D+5)^2y = 0$, i.e. $\{(D-1)^2+4\}^2y$, is

$$y = e^x\{(a+bx)\cos 2x+(c+dx)\sin 2x\},$$

so that the general solution of equation (iii) is

$$y = e^{-x}(L+Mx+Nx^2)+e^x\{(a+bx)\cos 2x+(c+dx)\sin 2x\}.$$

11.6. If $y = y_1$ is a solution of $L(D)y = 0$ and $y = \eta$ is a solution of $L(D)y = R$ then $y = y_1+\eta$ is a solution of $L(D)y = R$.

For $$L(D)(y_1+\eta) = L(D)y_1+L(D)\eta = R.$$

11.61. If $y = Y$ is the general solution of $L(D)y = 0$ and $y = \eta_1$ is any solution of $L(D)y = R$ then any other solution, $y = \eta_2$, of $L(D)y = R$ satisfies $\eta_2 = Y+\eta_1$ for a suitable choice of the constants in Y.

For $$L(D)(\eta_2-\eta_1) = L(D)\eta_2-L(D)\eta_1 = R-R = 0$$

so that $y = \eta_2-\eta_1$ is a solution of $L(D)y = 0$, and therefore, with a suitable choice of the constants in Y, $\eta_2-\eta_1 = Y$.

It follows that the equation L(D)y = R is completely solved if we know the general solution of L(D)y = 0 and any solution, however particular, of L(D)y = R itself. In relation to the solution of $L(D)y = R$ the solution of $L(D)y = 0$ is called the *Complementary Function* and the particular solution of $L(D)y = R$ is called a *Particular Integral*. The sum of the Complementary Function and a Particular Integral is called the *Complete Integral*. We have seen how to find the complementary function of the equation $L(D)y = R$; it remains to consider how a particular integral may be found.

11.62. If $y = \eta_1$ is a solution of $L(D)y = R_1$ and $y = \eta_2$ a solution of $L(D)y = R_2$ then $y = \eta_1+\eta_2$ is a solution of $L(D)y = R_1+R_2$.

For $$L(D)(\eta_1+\eta_2) = L(D)\eta_1+L(D)\eta_2 = R_1+R_2.$$

11.7. Particular integrals of the equation L(D)y = R when R takes one of the forms e^{px}, x^re^{px}, sin px, cos px, x^rsin px, x^rcos px, $e^{qx}x^r$sin px, $e^{qx}x^r$cos px, and x^r.

11.701. If $L(D)y = e^{px}$; since $L(D)e^{px} = e^{px}L(p)$, therefore, provided $L(p) \neq 0$, $L(D)\{e^{px}/L(p)\} = e^{px}$, and so a solution of $L(D)y = e^{px}$ is

$$y = e^{px}/L(p).$$

11.71. If $L(D) = (D-p)^k g(D)$, where $g(p) \neq 0$, then a solution of $L(D)y = e^{px}$ is
$$y = \frac{e^{px}}{g(p)} \frac{x^k}{k!}.$$

For $L(D+p)x^k = g(D+p)D^k x^k = g(D+p)k! = g(p)k!$

and so
$$e^{px}L(D+p)x^k = g(p)k!\,e^{px},$$
i.e.
$$L(D)e^{px}x^k = g(p)k!\,e^{px},$$
whence
$$L(D)\frac{e^{px}}{g(p)}\frac{x^k}{k!} = e^{px}.$$

11.72. If $y = y_1$ is a solution of $L(D)y = e^{px}$ then a solution of
$$L(D)y = x^r e^{px} \quad \text{is} \quad y = \frac{d^r y_1}{dp^r}.$$

Differentiate the equation $L(D)y_1 = e^{px}$ r times with respect to p and we find $L(D)\dfrac{d^r y_1}{dp^r} = x^r e^{px}$, proving that $y = \dfrac{d^r y_1}{dp^r}$ satisfies $L(D)y = x^r e^{px}$.

In particular, if the result of taking $p = 0$ in $d^r y_1/dp^r$, after differentiating, is denoted by y^* then $y = y^*$ is a solution of $L(D)y = x^r$.

We shall later give a more direct method of finding a solution of $L(D)y = x^r$.

Example. $(D^2-1)y = x(1+x)e^{2x}$.

The solution of $(D^2-1)y = e^{px}$ is
$$y = \frac{e^{px}}{p^2-1}$$
and so
$$\frac{dy}{dp} = \frac{xe^{px}}{p^2-1} - \frac{2pe^{px}}{(p^2-1)^2}$$
and
$$\frac{d^2y}{dp^2} = \frac{x^2e^{px}}{p^2-1} - \frac{4xpe^{px}}{(p^2-1)^2} + \frac{2(3p^2+1)e^{px}}{(p^2-1)^3}.$$
Taking $p = 2$ we see that a solution of $(D^2-1)y = xe^{2x}$ is
$$y = \frac{xe^{2x}}{3} - \frac{\text{[illegible]}}{9}e^{2x}$$
and a solution of $(D^2-1)y = x^2e^{2x}$ is
$$y = \frac{x^2e^{2x}}{3} - \frac{8}{9}xe^{2x} + \frac{26}{27}e^{2x}$$

and therefore a solution of $(D^2-1)y = x(1+x)e^{2x}$ is

$$y = \frac{x^2}{3}e^{2x} - \frac{5}{9}xe^{2x} + \frac{14}{27}e^{2x} \quad \text{(using § 11.62).}$$

11.8. To solve $L(D^2)y = \cos ax$.

Since $$L(D^2)\cos ax = L(-a^2)\cos ax$$

we have

$$L(D^2)\frac{\cos ax}{L(-a^2)} = \cos ax \quad \textit{provided } L(-a^2) \neq 0$$

so that the solution is

$$y = \frac{\cos ax}{L(-a^2)}.$$

Similarly the solution of $L(D^2)y = \sin ax$ is $y = \dfrac{\sin ax}{L(-a^2)}$.

11.81. If $L(D^2) = (D^2+a^2)^k g(D^2)$, where $g(-a^2) \neq 0$, then

$$\begin{aligned} L(D^2)x^k\cos(ax+\tfrac{1}{2}k\pi) &= g(D^2)(D^2+a^2)^k x^k \cos(ax+\tfrac{1}{2}n\pi) \\ &= g(D^2)(2a)^k k!\cos ax \\ &= g(-a^2)(2a)^k k!\cos ax. \end{aligned}$$

Hence the solution of

$$g(D^2)(D^2+a^2)^k y = \cos ax$$

is $$y = \frac{x^k}{k!}\frac{\cos(ax+\frac{1}{2}k\pi)}{(2a)^k g(-a^2)}.$$

11.82. If $L(D)$ is not a polynomial in D^2 then $L(D)$ may be expressed in the form $f(D^2)+Dg(D^2)$.

Then $$L(D)y = \cos ax$$

takes the form $$\{f(D^2)+Dg(D^2)\}y = \cos ax. \tag{i}$$

But

$$\begin{aligned} &\{f(D^2)+Dg(D^2)\}\{f(D^2)-Dg(D^2)\} \\ &\qquad = \{f(D^2)\}^2 - D^2\{g(D^2)\}^2 = h(D^2), \quad \text{say,} \end{aligned}$$

and so

$$\begin{aligned} &\{f(D^2)+Dg(D^2)\}\{f(D^2)-Dg(D^2)\}\cos ax \\ &\qquad = h(D^2)\cos ax = h(-a^2)\cos ax, \end{aligned}$$

showing that a solution of equation (i) is

$$y = \{f(D^2)-Dg(D^2)\}\cos ax/h(-a^2) \quad \text{provided } h(-a^2) \neq 0.$$

Whence $$y = \frac{f(-a^2)}{h(-a^2)}\cos ax - D\frac{g(-a^2)}{h(-a^2)}\cos ax,$$

i.e. $$\mathbf{y = \{f(-a^2)\cos ax + ag(-a^2)\sin ax\}/h(-a^2).}$$

Since $h(-a^2) = \{f(-a^2)\}^2 + a^2\{g(-a^2)\}^2$, $h(-a^2)$ can vanish only if both $f(-a^2)$ and $g(-a^2)$ are zero. If $f(-a^2) = g(-a^2) = 0$, let $(D^2+a^2)^k$ be the greatest power of D^2+a^2 which divides both $f(D^2)$ and $g(D^2)$ and let the quotients of $f(D^2)$, $g(D^2)$ by $(D^2+a^2)^k$ be $F(D^2)$ and $G(D^2)$ so that $F(-a^2)$ and $G(-a^2)$ do not vanish simultaneously. Further let $H(D^2) = \{F(D^2)\}^2 - D^2\{G(D^2)\}^2$. Then $L(D) = (D^2+a^2)^k\{F(D^2)+DG(D^2)\}$ and so, since

$$\begin{aligned}(D^2+a^2)^k\{F(D^2)+DG(D^2)\}\{F(D^2)-DG(D^2)\}x^k\cos(ax+\tfrac{3}{2}k\pi)\\ = H(D^2)(D^2+a^2)^k x^k\cos(ax+\tfrac{3}{2}k\pi) = H(D^2)(2a)^k k!\cos ax\\ = H(-a^2)(2a)^k k!\cos ax,\end{aligned}$$

therefore a solution of $L(D)y = \cos ax$ is

$$\begin{aligned}y &= \{F(D^2)-DG(D^2)\}\frac{x^k\cos(ax+\frac{3}{2}k\pi)}{k!(2a)^k H(-a^2)}\\ &= \left\{\frac{L(-D)}{(D^2+a^2)^k}\right\}\frac{x^k\cos(ax+\frac{3}{2}k\pi)}{k!(2a)^k H(-a^2)}.\end{aligned}$$

Alternatively, since

$$\begin{aligned}L(D)L(-D) &= \{f(D^2)\}^2 - D^2\{G(D^2)\}\\ &= (D^2+a^2)^{2k}H(D^2),\end{aligned}$$

therefore

$$\begin{aligned}L(D)L(-D)x^{2k}\cos(ax+3k\pi) &= H(D^2)(D^2+a^2)^{2k}x^{2k}\cos(ax+3k\pi)\\ &= H(D^2)(2a)^{2k}(2k)!\cos ax\\ &= H(-a^2)(2a)^{2k}(2k)!\cos ax,\end{aligned}$$

and so another solution of $L(D)y = \cos ax$ is

$$y = \frac{L(-D)x^{2k}\cos(ax+3k\pi)}{H(-a^2)(2a)^{2k}(2k)!}.$$

11.83. To solve $L(D)y = e^{px}\cos ax$.

We have $$e^{-px}L(D)y = \cos ax$$

and so $$L(D+p)ye^{-px} = \cos ax,$$

which is of the form $M(D)z = \cos ax$, and the solution is completed as in 11.82.

11.84. If $y = y_1$ is a solution of $L(D)y = e^{px}\cos ax$ then
$$L(D)y_1 = e^{px}\cos ax,$$
and so, differentiating r times with respect to p, we hav
$$L(D)\frac{d^r y_1}{dp^r} = x^r e^{px}\cos ax,$$
which shows that a solution of the equation
$$L(D)y = x^r e^{px}\cos ax$$
is
$$y = \frac{d^r y_1}{dp^r}$$
where $y = y_1$ is a solution of $L(D)y = e^{px}\cos ax$.

Similarly, if $y = y_1$ is a solution of $L(D)y = \cos(ax+\frac{1}{2}r\pi)$ ther $y = \frac{d^r y_1}{da^r}$ is a solution of $L(D)y = x^r\cos ax$.

EXAMPLES. A solution of $(D^2+2D-3)y = 3e^{2x}$ is
$$y = \frac{3e^{2x}}{2^2+2.2-3} = \cdot$$

A solution of $(D-2)^3(D^2+D+1)y = e^{2x}$ is
$$y = \frac{}{2^2+2+1}\,\frac{}{3!} = \frac{x^3e^{2x}}{42}$$

To find a solution of $(D^3+2D^2+3D+8)y = \cos 3x$.

Since $\qquad D^3+2D^2+3D+8 = 2D^2+8+D(D^2+3)$

we consider
$$\{(2D^2+8)+D(D^2+3)\}\{(2D^2+8)-D(D^2+3)\}\cos 3x$$
$$= \{(2D^2+8)^2-D^2(D^2+3)^2\}\cos 3x$$
$$= \{(-18+8)^2+9(-9+3)^2\}\cos 3x = 424\cos 3x,$$
which shows that a solution of the equation
$$\{(2D^2+8)+D(D^2+3)\}y = \cos 3x$$
is
$$y = \{(2D^2+8)-D(D^2+3)\}\frac{\cos 3x}{424}$$
$$= -\frac{10}{424}\cos 3x+\frac{6}{424}D\cos 3x$$
$$= -\frac{5}{212}\cos 3x-\frac{9}{212}\sin 3x$$

To find a solution of $(D^2-4D+5)^2(D^2-4D+6)y = e^{2x}\cos x$.

We have

$$e^{-2x}(D^2-4D+5)^2(D^2-4D+6)y = \cos x$$

and so $$(D^2+1)^2(D^2+2)ye^{-2x} = \cos x.$$

Now

$$(D^2+2)(D^2+1)^2x^2\cos(x+3\pi) = (D^2+2)2^2(2!)\cos x = 8\cos x$$

and so a solution of the equation is

$$ye^{-2x} = \frac{x^2}{8}\cos(x+3\pi) = -\frac{x^2}{8}\cos x,$$

i.e. $$y = -\tfrac{1}{8}e^{2x}x^2\cos x.$$

In particular cases it may be possible to make $L(D)$ a function of D^2 without multiplying by $L(-D)$, by using only some of the factors of $L(-D)$. For example, to solve $(D^2+4)^3(D+1)y = \cos 2x$ it suffices to consider $(D^2+4)^3(1-D^2)y = \cos 2x$; since

$$(D^2+4)^3x^3\cos(2x+\tfrac{3}{2}\pi) = 4^3(3!)\cos 2x$$

therefore

$$(1-D^2)(D^2+4)^3x^3\cos(2x+\tfrac{3}{2}\pi) = 4^3.3!(1+4)\cos 2x$$

so that a solution of $(D^2+4)^3(D+1)y = \cos 2x$ is

$$y = \frac{(1-D)x^3\cos(2x+\frac{3}{2}\pi)}{4^3.(3!).5} = -(\sin 2x)\frac{(x^3-3x^2)}{2^4.5!}+\frac{x^3\cos 2x}{2^3.5!}.$$

(Of course a solution of $(D^2+4)^3(1-D^2)y = \cos 2x$ may be written down directly by 11.81.)

11.9. The equation $L(D)y = R$ when $L(t)$ is a polynomial of the nth degree and R is a polynomial in x of the kth degree. We assume that $L(0) \neq 0$, for if $L(D) = D^pL^*(D)$, then integrating $D^pL^*(D)y = R$ repeatedly we find $L^*(D)y = S$, where $L^*(0) \neq 0$.

Suppose we can find a polynomial $M(t)$ which is of the kth degree and which is such that $L(t)M(t)-1$ contains no term of degree less than t^{k+1}, so that $L(t)M(t)-1$ is of the form $t^{k+1}N(t)$, then

$$\{L(D)M(D)-1\}R = N(D)D^{k+1}R = 0, \quad \text{since } R \text{ is of the } k\text{th degree},$$

and so $$L(D)M(D)R = R,$$

which shows that a solution of the equation $L(D)y = R$ is

$$y = M(D)R.$$

It remains to show how to find the appropriate polynomial $M(t)$.

The condition that $L(t)M(t)-1$ contains no term of degree less than t^{k+1} means that the coefficients of t, t^2, t^3,..., t^k in the product $L(t)M(t)$ are zero and the coefficient of t^0 is unity, so that we have $k+1$ conditions to determine the $k+1$ coefficients of the polynomial $M(t)$. Alternatively we can determine $M(t)$ as follows:

We have $$L(t)M(t) = 1+t^{k+1}N(t)$$

and so $$M(t) = \frac{1}{L(t)}+t^{k+1}\frac{N(t)}{L(t)}$$

Divide $L(t)$ into unity until we obtain a remainder $R(t)$ divisible by t^{k+1} and a quotient $Q(t)$ of the kth degree and let $R(t) = t^{k+1}S(t)$. Then

$$M(t) = Q(t)+t^{k+1}\frac{N(t)+S(t)}{L(t)}.$$

Since $t^{k+1}\dfrac{N(t)+S(t)}{L(t)}$ is equal to the polynomial $M(t)-Q(t)$ and since $L(t)$ is not divisible by t, therefore $L(t)$ must divide $N(t)+S(t)$ and so $t^{k+1}\dfrac{N(t)+S(t)}{L(t)}$ is either a polynomial of degree $(k+1)$ at *least*, or is zero. But $M(t)-Q(t)$ is of degree k at *most* and therefore $N(t)+S(t) = 0$ and $M(t) = Q(t)$, showing that the required polynomial $M(t)$ is obtained by dividing $L(t)$ (ordered by increasing powers of t) into unity until the quotient is of the kth degree.

11.91. To solve $L(D)y = Re^{px}$, where R is a polynomial, write

$$e^{-px}L(D)y = R,$$

whence $$L(D+p)ye^{-px} = R$$

and the solution proceeds as in 11.9.

Examples. I. To solve $(D+2)^2(D+4)^2y = e^{-3x}x^2(1+x^3)$.

We have $$e^{3x}(D+2)^2(D+4)^2y = x^2(1+x^3)$$

and so $$(D-1)^2(D+1)^2ye^{3x} = x^2+x^5,$$

i.e. $$(D^2-1)^2ye^{3x} = x^2+x^5.$$

Now $$\frac{1}{(1-t^2)^2} = 1+2t^2+3t^4+4t^6+...$$

and so a solution of the differential equation is given by

$$ye^{3x} = (1+2D^2+3D^4)(x^2+x^5)$$
$$= x^2+4+x^5+40x^3+360x = x^5+40x^3+x^2+360x+4,$$

i.e. $$y = e^{-3x}(x^5+40x^3+x^2+360x+4).$$

II. To solve $(D-5)^3(D-6)^2y = x^4e^{5x}$.

We have $$(D-1)^2D^3ye^{-5x} = x^4,$$
whence $$D^3ye^{-5x} = (1+2D+3D^2+4D^3+5D^4)x^4$$
$$= x^4+8x^3+36x^2+96x+120$$
and so
$$ye^{-5x} = \frac{x^7}{5.6.7}+\frac{8x^6}{6.5.4}+\frac{36x^5}{5.4.3}+\frac{96x^4}{4.3.2}+\frac{120x^3}{1.2.3},$$
i.e. $$y = \frac{x^3e^{5x}}{5.6.7}\{x^4+14x^3+126x^2+840x+4200\}.$$

11.92. *Variation of parameters. Lagrange's solution of*
$$L(D)y = R(x).$$

Let $y_1, y_2,..., y_n$ be a fundamental set of solutions of the equation $L(D)y = 0$, and define the set of functions $c_1(x), c_2(x),..., c_n(x)$ by the n equations
$$c_1 y_1+c_2 y_2+...+c_n y_n = 0,$$
$$c_1 Dy_1+c_2 Dy_2+...+c_n Dy_n = 0,$$
$$\cdots\cdots\cdots$$
$$c_1 D^{n-2}y_1+c_2 D^{n-2}y_2+...+c_n D^{n-2}y_n = 0,$$
$$c_1 D^{n-1}y_1+c_2 D^{n-1}y_2+...+c_n D^{n-1}y_n = R.$$
The equations are solvable for $c_1, c_2,..., c_n$ since the Wronskian of the set $y_1, y_2,..., y_n$ is not zero for any x.

Then $$y = y_1\int c_1\,dx+y_2\int c_2\,dx+...+y_n\int c_n\,dx$$
is the general solution of the equation $L(D)y = R$.

Proof. Write $C_r(x) = \int c_r(x)\,dx$, $1 \leqslant r \leqslant n$, so that
$$C_r'(x) = c_r(x),$$
then if $$y = y_1 C_1+y_2 C_2+...+y_n C_n$$
it follows that
$$Dy = C_1 Dy_1+C_2 Dy_2+...+C_n Dy_n,$$
$$\text{since } \sum y_r DC_r = \sum y_r c_r = 0,$$
$$D^2y = C_1 D^2y_1+C_2 D^2y_2+...+C_n D^2y_n,$$
$$\text{since } \sum c_r Dy_r = 0,$$

$$D^{n-1}y = C_1 D^{n-1}y_1+...+C_n D^{n-1}y_n,$$
$$\text{since } \sum c_r D^{n-2}y_r = 0,$$

and $D^n y = C_1 D^n y_1 + \ldots + C_n D^n y_n + R$, since $\sum c_r D^{n-1} y_r = R$, whence $L(D)y = C_1 L(D)y_1 + \ldots + C_n L(D)y_n + R = R$.

EXAMPLE. To solve $(D^2+1)y = \sec^2 x$.

The fundamental set of solutions of $(D^2+1)y = 0$ is $\sin x$, $\cos x$.

The solution of the equations

$$c_1 \cos x + c_2 \sin x = 0, \qquad -c_1 \sin x + c_2 \cos x = \sec^2 x$$

is $c_1 = -\sec x \tan x$, $c_2 = \sec x$. Hence the general solution is

$$y = \sin x \int \sec x \, dx - \cos x \int \sec x \tan x \, dx$$
$$= A \sin x + B \cos x + \sin x \log(\sec x + \tan x) - 1.$$

11.93. Simultaneous differential equations

11.931. *Routh's method.* The general solution of the system of equations

$$P_n u + Q_n v + R_n w = 0, \quad n = 1, 2, 3,$$

where u, v, w are functions of x and P_n, Q_n, R_n are polynomials in D_x, $n = 1, 2, 3$, is $u = \lambda_1 \zeta$, $v = \mu_1 \zeta$, $w = \nu_1 \zeta$, where λ_n, μ_n, ν_n are the cofactors of P_n, Q_n, R_n in the determinant Δ with nth row $P_n Q_n R_n$, $n = 1, 2, 3$, and ζ is the general solution of the equation $\Delta\zeta = 0$.

For

$$P_n u + Q_n v + R_n w = (P_n \lambda_1 + Q_n \mu_1 + R_n \nu_1)\zeta = 0$$
$$\text{for } all\ \zeta \text{ if } n = 2 \text{ or } 3$$
$$= \Delta\zeta \quad \text{if } n = 1,$$

and so all three equations are satisfied if $\Delta\zeta = 0$.

11.932. A particular integral of the system of equations

$$P_n u + Q_n v + R_n w = X_n, \quad n = 1, 2, 3$$

is

$$u = \lambda_1 \xi + \lambda_2 \eta + \lambda_3 \zeta, \quad v = \mu_1 \xi + \mu_2 \eta + \mu_3 \zeta, \quad w = \nu_1 \xi + \nu_2 \eta + \nu_3 \zeta,$$

where ξ, η, ζ are any particular integrals of the equations $\Delta\xi = X$ $\Delta\eta = X_2$, and $\Delta\zeta = X_3$ respectively.

For

$$P_1 u + Q_1 v + R_1 w = (P_1 \lambda_1 + Q_1 \mu_1 + R_1 \nu_1)\xi +$$
$$+ (P_1 \lambda_2 + Q_1 \mu_2 + R_1 \nu_2)\eta + (P_1 \lambda_3 + Q_1 \mu_3 + R_1 \nu_3)\zeta = \Delta\xi,$$

and similarly

$$P_2 u + Q_2 v + R_2 w = \Delta\eta, \qquad P_3 u + Q_3 v + R_3 w = \Delta\zeta.$$

11.933. If $u = U$, $v = V$, $w = W$ is the general solution of the equations $L_n \equiv P_n u + Q_n v + R_n w = 0$, $n = 1, 2, 3$, and $u = u_0$, $v = v_0$, $w = w_0$ is a particular integral of the system $L_n = X_n$, $n = 1, 2, 3$, then $u = U + u_0$, $v = V + v_0$, $w = W + w_0$ is the complete solution of the system $L_n = X_n$, $n = 1, 2, 3$.

Let $u = u^*$, $v = v^*$, $w = w^*$ be any solution of the equations $L_n = X_n$; then $u = u^* - u_0$, $v = v^* - v_0$, $w = w^* - w_0$ are solutions of $L_n = 0$, for

$$P_n(u^* - u_0) + Q_n(v^* - v_0) + R_n(w^* - w_0) = X_n - X_n = 0.$$

Hence $u^* - u_0 = U$ for some values of the constants in U, etc., and so $u = U + u_0$, $v = V + v_0$, $w = W + w_0$ is the complete solution.

EXAMPLE. Solve the system of equations

$$Du + v + (D+1)w = 0, \qquad (D+1)u + Dv + w = 0,$$
$$u + (D+1)v + Dw = 0.$$

We have $\Delta = 2(D^3+1)$ and so $(D^3+1)\zeta = 0$ whence

$$\zeta = Ae^{-x} + e^{\frac{1}{2}x}\left\{B\cos\frac{\sqrt{3}}{2}x + C\sin\frac{\sqrt{3}}{2}x\right\}.$$

Therefore

$$u = (D^2 - D - 1)\zeta = Ae^{-x} - 2e^{\frac{1}{2}x}\left\{B\cos\frac{\sqrt{3}}{2}x + C\sin\frac{\sqrt{3}}{2}x\right\},$$

$$v = (1 - D - D^2)\zeta$$
$$= Ae^{-x} + e^{\frac{1}{2}x}\left\{(B - C\sqrt{3})\cos\frac{\sqrt{3}}{2}x + (C + B\sqrt{3})\sin\frac{\sqrt{3}}{2}x\right\},$$

$$w = (D^2 + D + 1)\zeta$$
$$= Ae^{-x} + e^{\frac{1}{2}x}\left\{(B + C\sqrt{3})\cos\frac{\sqrt{3}}{2}x + (C - B\sqrt{3})\sin\frac{\sqrt{3}}{2}x\right\}.$$

XII

MEAN-VALUE THEOREMS

THE DERIVATIVE ATTAINS THE MEAN SLOPE. ROLLE'S THEOREM. THE CAUCHY FORMULA. THE GENERALIZED CAUCHY FORMULA. MEAN-VALUE THEOREMS FOR INTEGRALS

12. As in § 3.6 we shall denote $\{f(b)-f(a)\}/(b-a)$ by $\mu(a, b)$.

12.1. If X lies in $[a, b]$, and if $\mu(a, X) > \mu(a, b)$, then

$$\mu(X, b) < \mu(a, b).$$

For if $\quad \{f(X)-f(a)\}/(X-a) > \{f(b)-f(a)\}/(b-a)$

then $\quad f(X) > f(a)+\{f(b)-f(a)\}\{(X-a)/(b-a)\}$

and so

$$\begin{aligned} f(b)-f(X) &< \{f(b)-f(a)\}-\{f(b)-f(a)\}\{(X-a)/(b-a)\} \\ &= \{f(b)-f(a)\}\{(b-a)-(X-a)\}/(b-a) \\ &= \{f(b)-f(a)\}(b-X)/(b-a), \end{aligned}$$

whence $\quad \mu(X, b) < \mu(a, b).$

12.11. Similarly, if $\mu(a, X) < \mu(a, b)$ then $\mu(X, b) > \mu(a, b)$.

12.2. If $\{f(x)-f(a)\}/(x-a)$ is constant in (a, b) then $f'(x) = \mu(a, b)$ for any x in (a, b).

For $\{f(x)-f(a)\}/(x-a) = \mu(a, b)$ and so

$$f(x) = f(a)+\mu(a, b).(x-a),$$

whence $f'(x) = \mu(a, b)$.

12.21. If $\{f(x)-f(a)\}/(x-a)$ is not constant in (a, b), then we can determine a point c in $[a, b]$ such that $f'(c) = \mu(a, b)$.

For if $\mu(a, x)$ is not constant, there is an X in (a, b) such that $\mu(a, X)$ is different from $\mu(a, b)$. Suppose $\mu(a, X) > \mu(a, b)$, so that, by 12.1, $\mu(X, b) < \mu(a, b)$.

Now by 3.61 there is a point c_1 in (a, X) such that

$$f'(c_1) \geqslant \mu(a, X)$$

and a c_2 in (X, b) such that $f'(c_2) \leqslant \mu(X, b)$, and therefore

$$f'(c_1) > \mu(a, b) > f'(c_2).$$

Since $f'(x)$ is continuous (3.02), $f'(x)$ takes every value between

$f'(c_1)$ and $f'(c_2)$, and in particular there is, accordingly, a point c in the open interval $[c_1, c_2]$ such that

$$f'(c) = \mu(a, b).$$

Since c lies in $[c_1, c_2]$, therefore c lies in $[a, b]$.

Similarly, if $\mu(a, X) < \mu(a, b)$ there is a point c in $[a, b]$ such that $f'(c) = \mu(a, b)$.

Theorems 12.2 and 12.21 together give:

12.22. If f(x) is differentiable in (a, b) then there is a point c such that f′(c) = {f(b)−f(a)}/(b−a) and a < c < b.

Theorem 12.22 is known as the *mean-value theorem.*

12.23. It is important to observe that in the mean-value theorem we prove the existence of a point c where the derivative equals the mean slope in (a, b), and ensure that this point lies between a and b, and does not coincide with an end-point of the interval. The existence of a point in the *closed* interval (a, b) where the derivative equals the mean slope—a result of relatively little importance—is of course a consequence of 12.22 but may be established more simply as follows: We know directly from 3.61 that there are points c_1, c_2 in (a, b) such that $f'(c_1) \leqslant \mu(a, b)$ and $f'(c_2) \geqslant \mu(a, b)$; if equality occurs at either place the result is established, and if there is no equality, then $f'(c_1) < \mu(a, b) < f'(c_2)$ and Theorem 12.21 follows.

If $f(a) = f(b)$ then $\mu(a, b) = 0$ and from the *mean-value theorem* we deduce *Rolle's theorem* that

12.3. If f(x) is differentiable in (a, b), and if f(a) = f(b), then there is a point c such that f′(c) = 0 and a < c < b.

Although apparently just a special case of the mean-value theorem, Rolle's theorem is in fact as general as the mean-value theorem; we shall show this by establishing Rolle's theorem independently of 12.22 and then deriving the mean-value theorem from it.

12.4. If $f(x)$ is constant in (a, b) then $f'(x) = 0$ throughout (a, b).

12.41. If $f(a) = f(b)$ and $f(x)$ is not constant in (a, b), then there is a point c such that $f'(c) = 0$ and $a < c < b$.

For if $f(x)$ is not constant there is a point X in (a, b) where $f(X)$ differs from $f(a)$; suppose that $f(X) > f(a)$, then since $f(b) = f(a)$ it follows that $f(X) > f(b)$.

By 3.61 there is a point c_1 in (a, X) such that

$$f'(c_1) \geqslant \{f(X)-f(a)\}/(X-a) > 0,$$

and a point c_2 in (X, b) such that

$$f'(c_2) \leqslant \{f(b)-f(X)\}/(b-X) < 0;$$

hence since $f'(x)$ is continuous, positive at c_1 and negative at c_2, therefore there is a point c *between* c_1 and c_2 such that $f'(c) = 0$. Similarly, if $f(X) < f(a)$, there is a point c in $[a, b]$ where the derivative vanishes.

12.42. Theorems 12.4 and 12.41 together give Rolle's theorem. Consider now the function $\phi(x) = f(x)-Ax$; $\phi(x)$ is differentiable and $\phi(a) = \phi(b)$, provided $f(a)-Aa = f(b)-Ab$, i.e. provided $A = \mu(a, b)$. Thus $\phi(x)$ satisfies the conditions of Rolle's theorem and so there is a point c in $[a, b]$ such that $\phi'(c) = 0$; but

$$\phi'(x) = f'(x)-\mu(a, b)$$

and so $f'(c) = \mu(a, b)$, which proves 12.22.

12.5. If $f(x)$ and $g(x)$ are differentiable in (a, b) and if $g'(x) \neq 0$ in $[a, b]$ then we can determine a point c in $[a, b]$ such that

12.51. $$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)} \quad \text{(Cauchy)}$$

Consider the function $L(x) = \{f(b)-f(a)\}g(x)-\{g(b)-g(a)\}f(x)$; $L(x)$ is differentiable, and $L(a) = L(b) = f(b)g(a)-f(a)g(b)$. Hence, by 12.3, there is a point c in $[a, b]$ such that $L'(c) = 0$, i.e.

$$\{f(b)-f(a)\}g'(c) = \{g(b)-g(a)\}f'(c).$$

Since $g'(x) \neq 0$ in $[a, b]$, therefore $g'(c) \neq 0$, and $g(b) \neq g(a)$, whence equation 12.51 follows.

The interest in 12.51 lies in the fact that the variable on the right-hand side has the same value c in both numerator and denominator, an identity which cannot be ensured by the application of the Mean-Value Theorem to $f(x)$ and $g(x)$ separately.

The Cauchy formula (12.51) is really an example of the application of the Mean-Value Theorem to a composite function.

For if $G(x)$ is a function differentiable in an interval (α, β), with a non-vanishing derivative in $[\alpha, \beta]$, then, since the derivative of $f(G(x))$ is $f'(G(x))G'(x)$, by the Mean-Value Theorem there is a point γ in (α, β) such that

$$\frac{f(G(\beta))-f(G(\alpha))}{\beta-\alpha} = f'(G(\gamma))G'(\gamma). \tag{i}$$

Since $G'(x) \neq 0$ in (α, β) and $G'(x)$ is necessarily continuous, therefore $G'(x)$ is of constant sign, and so $G(x)$ has a unique differentiable inverse $g(x)$, say, such that $G'(x) = 1/g'(G(x))$. Write a, b, c for $G(\alpha)$, $G(\beta)$, $G(\gamma)$ so that $g(a) = \alpha$, $g(b) = \beta$, and equation (i) becomes

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

12.52. The functions $f(x)$, $g(x)$ are differentiable n times in the interval (α, b),

$$f^r(\alpha) = g^r(\alpha) = 0, \quad 1 \leqslant r < n,$$

and $$|g^r(x)| > 0 \quad \text{in} \quad [\alpha, b], \quad 1 \leqslant r \leqslant n.$$

Then if $\alpha \leqslant a < b$, we can find c such that

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f^n(c)}{g^n(c)}.$$

For by Cauchy's formula we can find $c_1, c_2, \ldots, c_{n-1}, c$ in turn, such that

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c_1)}{g'(c_1)}$$

and $$\frac{f'(c_1)}{g'(c_1)} = \frac{f'(c_1)-f'(\alpha)}{g'(c_1)-g'(\alpha)} = \frac{f^2(c_2)}{g^2(c_2)} = \ldots = \frac{f^{n-1}(c_{n-1})}{g^{n-1}(c_{n-1})}$$

$$= \frac{f^{n-1}(c_{n-1})-f^{n-1}(\alpha)}{g^{n-1}(c_{n-1})-g^{n-1}(\alpha)} = \frac{f^n(c)}{g^n(c)},$$

which completes the proof.

Note that c_1 lies in $[a, b]$, so that $c_1 > \alpha$, and since c_{r+1} lies in $[\alpha, c_r]$ therefore $c_r > \alpha$ implies that $c_{r+1} > \alpha$; hence $c_r > \alpha$ for all r, and therefore $|g^r(x)| > 0$ in each of the intervals $[\alpha, c_n]$, which justifies the applications of Cauchy's formula made in the proof. The conditions $\alpha \leqslant a < b$ may be replaced by $a < b \leqslant \alpha$; the essential point is that the derivatives $g^r(x)$ should not vanish *between* a and b.

12.53. The functions $f(x)$, $g(x)$ are differentiable n times in (a, b), and $|g^n(x)| > 0$ in (a, b).

If $f(x)$, $g(x)$ vanish simultaneously at n distinct points between a and b then there is a point $c = c(x)$ between a and b such that

$$\frac{f(x)}{g(x)} = \frac{f^n(c)}{g^n(c)}$$

at all points of (a, b) where $g(x) \neq 0$.

Proof. Let

$$H(t) = f(x)g(t) - g(x)f(t), \quad a < x < b,\ g(x) \neq 0,$$

then $H(t)$ is differentiable n times in (a, b) and $H(t)$ vanishes a $n+1$ distinct points in (a, b), say at $t_1, t_2,..., t_n$, and x. Hence b Rolle's theorem we can find n points $u_1, u_2,..., u_{n-1}, u_n$ in $(a, b$ such that

$$H'(u_r) = 0, \quad r = 1, 2,..., n-1, n.$$

It follows that there are (at least) $n-1$ points in $[a, b]$ where $H^2(t$ vanishes, $n-2$ points where $H^3(t)$ vanishes, and so on up to 2 point where $H^{n-1}(t)$ vanishes, and so finally we reach a point c in $[a, b$ where $H^n(t) = 0$. But $H^n(t) = f(x)g^n(t) - g(x)f^n(t)$ and so

$$f(x)g^n(c) = g(x)f^n(c)$$

or

$$\frac{f(x)}{g(x)} = \frac{f^n(c)}{g^n(c)}$$

provided $g(x) \neq 0$ (for $g^n(c) \neq 0$ by hypothesis).

As an application of 12.53 we prove that if $\phi(x)$, $\psi(x)$ are twic differentiable in (y, z) and if x is a point in this interval then ther is a point c in $[y, z]$ such that

12.54 $$\frac{(y-z)\phi(x)+(z-x)\phi(y)+(x-y)\phi(z)}{(y-z)\psi(x)+(z-x)\psi(y)+(x-y)\psi(z)} = \frac{\phi^2(c)}{\psi^2(c)}.$$

For i

$$f(x) = (y-z)\phi(x)+(z-x)\phi(y)+(x-y)\phi(z),$$
$$g(x) = (y-z)\psi(x)+(z-x)\psi(y)+(x-y)\psi(z)$$

then both $f(x)$ and $g(x)$ vanish at the end-points of the interval (y, z).

12.541. Taking $\psi(x) = (x-y)(z-x)$ in 12.54 we obtain the *interpolation* formula

$$\phi(x) = \frac{x-y}{z-y}\phi(z) + \frac{z-x}{z-y}\phi(y) - \tfrac{1}{2}(x-y)(z-x)\phi^2(c).$$

12.542. In determinant notation formula 12.54 may be written

$$\psi^2(c)\begin{vmatrix} \phi(x) & x & 1 \\ \phi(y) & y & 1 \\ \phi(z) & z & 1 \end{vmatrix} = \phi^2(c)\begin{vmatrix} \psi(x) & x & 1 \\ \psi(y) & y & 1 \\ \psi(z) & z & 1 \end{vmatrix},$$

which immediately suggests the extension to four or more variables.

12.55. If $\phi(x)$, $\psi(x)$ are differentiable three times in an interval which contains the four points x, y, z, w then there is a point c in this interval such that

$$\psi^3(c)\begin{vmatrix} \phi(x) & x^2 & x & 1 \\ \phi(y) & y^2 & y & 1 \\ \phi(z) & z^2 & z & 1 \\ \phi(w) & w^2 & w & 1 \end{vmatrix} = \phi^3(c)\begin{vmatrix} \psi(x) & x^2 & x & 1 \\ \psi(y) & y^2 & y & 1 \\ \psi(z) & z^2 & z & 1 \\ \psi(w) & w^2 & w & 1 \end{vmatrix}.$$

This is an immediate consequence of Theorem 12.53, with $n = 3$, since each determinant is a function of x which vanishes at the three points y, z, w.

An alternative method of proof is to observe that the determinant

$$H(t) = \begin{matrix} \phi(t) & \psi(t) & t^2 & t \\ \phi(x) & \psi(x) & x^2 & x \\ \phi(y) & \psi(y) & y^2 & y \\ \phi(z) & \psi(z) & z^2 & z \\ \phi(w) & \psi(w) & w^2 & w \end{matrix}$$

is a function of t which vanishes at the *four* points x, y, z, and w, so that $H^3(t)$ vanishes at some point c. Only the first line of the determinant varies with t, so that the third derivative of the determinant is obtained by differentiating the elements of the first row three times, which leads immediately to the required formula.

12.6. Integral mean-value theorems

If we apply the mean-value theorem to the function

$$f(x) = \int_a^x g(t)\,dt,$$

since $f'(x) = g(x)$, we have the mean-value theorem for integrals

$$\int_a^b g(x)\,dx = (b-a)g(c).$$

We have already obtained this result in 9.13; the present proof has the advantage of not assuming a knowledge of the least and greatest values of $g(x)$ in (a, b).

12.61. Applying 12.5 to the functions

$$F(x) = \int_a^x f(t)g(t)\,dt, \qquad G(x) = \int_a^x g(t)\,dt$$

with $g(x) > 0$ in $[a, b]$, so that $G(x)$ is increasing in (a, b), we have

$$\frac{F(b)-F(a)}{G(b)-G(a)} = \frac{F'(c)}{G'(c)}$$

i.e

$$\left\{\int_a^b f(x)g(x)\,dx\right\}\Big/\left\{\int_a^b g(x)\,dx\right\} = f(c)g(c)/g(c) = f(c), \quad \text{since } g(c) \neq 0,$$

whence

12.611. $$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx.$$

12.611 is known as the *second mean-value theorem for integrals.*

12.62. In 12.61 the condition '$g(x) > 0$' in $[a, b]$ may be replaced by '$g(x) < 0$' in $[a, b]$, for if $g(x) < 0$ then $-g(x) > 0$ and so by 12.611

$$\int^b f(x)\{-g(x)\}\,dx = f(c)\int^b \{-g(x)\}\,dx,$$

whence $$\int^b f(x)g(x)\,dx = f(c)\int^b g(x)\,dx.$$

12.7. If $f(x)$ and $g(x)$ are continuous in (a, b) and if $f(x)$ is monotonic (*increasing* or *decreasing*), and $g(x) > 0$, in $[a, b]$ then there is a point c in $[a, b]$ such that

12.71. $$\int_a^b f(x)g(x)\,dx = f(a)\int_a^c g(x)\,dx + f(b)\int_c^b g(x)\,dx.$$

Let $G(x)$ denote the function $f(a)\int_a^x g(x)\,dx + f(b)\int_x^b g(x)\,dx$, and k the integral $\int_a^b g(x)\,dx$.

Then $G(a) = kf(b)$ and $G(b) = kf(a)$, and by 12.611 there is a point X in $[a, b]$ such that

$$\int^b f(x)g(x)\,dx = f(X)\int^b g(x)\,dx = kf(X).$$

Since X lies in $[a, b]$ and $f(x)$ is monotonic, therefore $kf(X)$ lies between $kf(a)$ and $kf(b)$, i.e. $kf(X)$ lies between $G(a)$ and $G(b)$; but $\int_a^b f(x)g(x)\,dx = kf(X)$, and so $\int_a^b f(x)g(x)\,dx$ lies between $G(a)$ and $G(b)$. Therefore, as $G(x)$ is continuous in (a, b), we can determine a point c in $[a, b]$ such that $G(c) = \int_a^b f(x)g(x)\,dx$, which proves 12.71.

12.72. $f(x)$ is monotonic and continuous, and does not change sign in (a, b), and $g(x)$ is continuous in (a, b) and greater than zero in $[a, b]$.

If $f(x)$ is positive and decreasing (or negative and increasing) then there is a point c in $[a, b]$ such that

12.73. $$\int_a^b f(x)g(x)\,dx = f(a)\int_a^c g(x)\,dx$$

and if $f(x)$ is positive and increasing (or negative and decreasing) then there is a point c in $[a, b]$ such that

12.74. $$\int_a^b f(x)g(x)\,dx = f(b)\int_c^b g(x)\,dx.$$

Proof of 12.73. If $f(x)$ is positive and decreasing then

$$\{f(a)-f(x)\}g(x) \geqslant 0, \quad x \geqslant a,$$

and so $\int_a^b \{f(a)-f(x)\}g(x)\,dx > 0$, whence

$$0 \leqslant \int_a^b f(x)g(x)\,dx < f(a)\int_a^b g(x)\,dx$$

and so $\int_a^b f(x)g(x)\,dx$ lies between the values of $f(a)\int_a^x g(x)\,dx$ for $x = a$ and $x = b$. Since $\int_a^x g(x)\,dx$ is continuous there is, therefore, a point c in $[a, b]$ such that

$$f(a)\int^c g(x)\,dx = \int^b f(x)g(x)\,dx,$$

i.e. $$\int^b f(x)g(x)\,dx = f(a)\int^c g(x)\,dx.$$

If $f(x)$ is negative and decreasing then $-f(x)$ is positive an increasing and so

$$\int_a^b \{-f(x)\}g(x)\,dx = -f(a)\int_a^c g(x)\,dx$$

whence
$$\int_a^b f(x)g(x)\,dx = f(a)\int_a^c g(x)\,dx.$$

Proof of 12.74. If $f(x)$ is positive and increasing the

$$\{f(b)-f(x)\}g(x) \geqslant 0 \quad \text{for} \quad x \leqslant b$$

and so $\int_a^b f(x)g(x)\,dx$ lies between zero and $f(b)\int_a^b g(x)\,dx$, whence the proof is completed as in 12.73.

12.75. We may interchange $f(a)$ and $f(b)$ in 12.71; the only change in the proof lies in taking $G(x) = f(b)\int_a^x g(x)\,dx + f(a)\int_x^b g(x)\,dx$.

In 12.73 we may replace the right-hand side by $f(a)\int_c^b g(x)\,dx$, for $\int_a^b f(x)g(x)\,dx$ lies between $f(a)\int_b^b g(x)\,dx$ and $f(a)\int_a^b g(x)\,dx$ and so $\left(\int_a^b f(x)g(x)\,dx\right)\Big/f(a)$ is a value of $\int_x^b g(x)\,dx$ for an x between a and b.

Similarly, in 12.74 we may replace the right-hand side by $f(b)\int_a^c g(x)\,dx$. The point to notice is that when $f(x)$ is positive and decreasing we take $f(a)$ outside the integral, and when $f(x)$ is positive and increasing we take $f(b)$ outside.

12.76. In Theorem 12.72 the condition that $g(x)$ is greater than zero in $[a, b]$ can be relaxed. We shall show that formulae 12.73 and 12.74 hold even when $g(x)$ changes sign in $[a, b]$, *provided that $g(x)$ vanishes only a finite number of times in* $[a, b]$. It suffices to consider the case when, for instance, $g(x)$ vanishes twice in $[a, b]$, the proof applying, unchanged, to any other case. We prove in fact:

12.761. If $f(x)$ is positive, increasing, and continuous in (a, b), and if $g(x)$ is continuous in (a, b) and vanishes only for $x = \alpha$ and $x = \beta$ in $[a, b]$, $\alpha < \beta$, then there is a point c in $[a, b]$ such that

$$\int^{b} f(x)g(x)\,dx = f(b)\int^{b} g(x)\,dx.$$

Proof. Since $g(x)$ is continuous and non-zero in $[a, \alpha]$, $[\alpha, \beta]$, and $[\beta, b]$ therefore $g(x)$ is of constant sign in these intervals, whence by 12.61 and 12.62 we can find c_1, c_2, c_3 in $[a, \alpha]$, $[\alpha, \beta]$, $[\beta, b]$ respectively such that

$$\int^{b} f(x)g(x)\,dx$$

$$= \int_{a}^{\alpha} f(x)g(x)\,dx + \int_{\alpha}^{\beta} f(x)g(x)\,dx + \int_{\beta}^{b} f(x)g(x)\,dx$$

$$= f(c_1)\int_{a}^{\alpha} g(x)\,dx + f(c_2)\int_{\alpha}^{\beta} g(x)\,dx + f(c_3)\int_{\beta}^{b} g(x)\,dx$$

$$= f(c_1)\left\{\int_{a}^{b} - \int_{\alpha}^{b} g(x)\,dx\right\} + f(c_2)\left\{\int_{\alpha}^{b} - \int_{\beta}^{b} g(x)\,dx\right\} + f(c_3)\int_{\beta}^{b} g(x)\,dx$$

$$= f(c_1)\int_{a}^{b} g(x)\,dx + \{f(c_2)-f(c_1)\}\int_{\alpha}^{b} g(x)\,dx +$$
$$+\{f(c_3)-f(c_2)\}\int_{\beta}^{b} g(x)\,dx$$

$$= f(c_1)\int_{a}^{b} g(x)\,dx + \{f(c_2)-f(c_1)\}\int_{\alpha}^{b} g(x)\,dx +$$
$$+\{f(c_3)-f(c_2)\}\int_{\beta}^{b} g(x)\,dx + \{f(b)-f(c_3)\}\int_{b}^{b} g(x)\,dx,$$

$$\text{since } \int_{b}^{b} g(x)\,dx = 0;$$

hence if l and g are the least and greatest of the four numbers

$$\int_{a}^{b} g(x)\,dx,\quad \int_{\alpha}^{b} g(x)\,dx,\quad \int_{\beta}^{b} g(x)\,dx,\quad \text{and}\quad \int_{b}^{b} g(x)\,dx,$$

since

$$f(c_1),\quad \{f(c_2)-f(c_1)\},\quad f(c_3)-f(c_2),\quad \text{and}\quad \{f(b)-f(c_3)\}$$

are positive, and

$$f(c_1)+\{f(c_2)-f(c_1)\}+\{f(c_3)-f(c_2)\}+\{f(b)-f(c_3)\} = f(b$$

we ha
$$lf(b) \leqslant \int_a^b f(x)g(x)\,dx \leqslant gf(b).$$

Thus $\int_a^b f(x)g(x)\,dx$ lies between two values of the continuou function $f(b)\int_t^b g(x)\,dx$, $a < t < b$, and so there is a value of t i $[a, b]$, c (say), such that

$$\int_a^b f(x)g(x)\,dx = f(b)\int_c^b g(x)\,dx$$

12.762. If $f(x)$ is positive, decreasing, and continuous in (a, b) an if $g(x)$ is continuous in (a, b) and vanishes only for $x = \alpha$, $x = \beta$ in $[a, b]$, $\alpha < \beta$, then there is a point c in $[a, b]$ such that

$$\int_a^b f(x)g(x)\,dx = f(a)\int_a^c g(x)\,dx$$

For

$$\int_a^b f(x)g(x)\,dx = f(c_1)\int_a^\alpha + f(c_2)\int_\alpha^\beta + f(c_3)\int_\beta^b g(x)\,dx$$

$$= \{f(a)-f(c_1)\}\int_a^a + \{f(c_1)-f(c_2)\}\int_a^\alpha + \{f(c_2)-f(c_3)\}\int_a^\beta + f(c_3)\int_a^b$$

and so $\int_a^b f(x)g(x)\,dx$ lies between $lf(a)$ and $gf(a)$, where l and g are the least and greatest of the four numbers $\int_a^t g(x)\,dx$, $t = a, \alpha, \beta, b$, whence the proof is completed as above.

XIII

THE TAYLOR SERIES

STANDARD AND GENERAL REMAINDERS. GENERALIZATION OF THE LIMIT CONCEPT. CONTACT OF PLANE CURVES. APPROXIMATION TO THE ROOT OF AN EQUATION

13. In § 7.4 we proved that *if* the difference $f(X)-f(x)$ can be expanded in powers of $X-x$ in a convergent series

$$(X-x)a_1(x)+\frac{(X-x)^2}{2!}a_2(x)+\frac{(X-x)^3}{3!}a_3(x)+\ldots+\frac{(X-x)^n}{n!}a_n(x)+\ldots,$$

then, for all values of n, $a_n(x)$ is equal to $f^n(x)$, the nth derivative of $f(x)$; in other words, the expansion

13.1. $$f(x)+(X-x)f'(x)+\frac{(X-x)^2}{2!}f''(x)+\frac{(X-x)^3}{3!}f'''(x)+\ldots+$$
$$+\frac{(X-x)^n}{n!}f^n(x)+\ldots$$

is the only possible expansion of $f(X)$ in powers of $X-x$.

The series 13.1 is called the *Taylor series* associated with the function $f(X)$, in the interval (x, X).

We have seen that the equality of $f(X)$ and its Taylor series is assured if we *know* that $f(X)$ can be expanded in a series, but this condition is too indirect to be of much practical value, and we shall now consider what simple restrictions imposed upon $f(x)$ directly suffice to ensure the equality of $f(X)$ with its Taylor series.

We shall first consider the difference between $f(X)$ and the sum to n terms of the Taylor series 13.1; this difference is known as the *remainder* (after n terms) in the Taylor series. We shall denote the sum to n terms of 13.1 by $T(x)$ (or by $T_n(x)$ when we require the 'n' explicitly), and make the following initial assumption about $f(x)$:

13.2. $f(x)$ is differentiable n times in the interval (x, X).

13.21. Since

$$T_{r+1}(x) = T_r(x)+\frac{(X-x)^r}{r!}f^r(x) \quad \text{and} \quad T_1(x) = f(x),$$

therefoı $$T_n(X) = f(X).$$

an

$$T'_{r+1}(x) = T'_r(x) - \frac{(X-x)^{r-1}}{(r-1)!} f^r(x) + \frac{(X-x)^r}{r!} f^{r+1}(x)$$

$$T'_1(x) = f'(x);$$

hence $T'_2(x) = (X-x)f^2(x)$, $\quad T'_3(x) = \dfrac{(X-x)^2}{2!} f^3(x)$,

and so on up to $\quad T'_n(x) = \dfrac{(X-x)^{n-1}}{(n-1)!} f^n(x).$

Since the function $T(x)$ satisfies the conditions of the mean-value theorem, 12.22, we can determine a point c in (x, X) such that

$$T(X) - T(x) = (X-x)T'(c).$$

Accordingly

$$f(X) - T(x) = T(X) - T(x) = (X-x)\frac{(X-c)^{n-1}}{(n-1)!} f^n(c).$$

Thus

13.3. $$f(X) = f(x) + (X-x)f'(x) + \ldots + \frac{(X-x)^{n-1}}{(n-1)!} f^{n-1}(x) + \\ + (X-x)\frac{(X-c)^{n-1}}{(n-1)!} f^n(c).$$

If we write $X-x = h$ and $(c-x)/(X-x) = \theta$, so that $c = x+\theta h$ and $X-c = h(1-\theta)$, then equation 13.3 takes the form

13.31. $$\mathbf{f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \ldots + \frac{h^{n-1}}{(n-1)!} f^{n-1}(x) + \\ + h^n \frac{(1-\theta)^{n-1}}{(n-1)!} f^n(x+\theta h).}$$

The remainder $\mathbf{h^n \dfrac{(1-\theta)^{n-1}}{(n-1)!} f^n(x+\theta h)}$ is known as *Cauchy's* form.

13.311. We obtain a second form of the remainder by applying Theorem 12.5 to the functions $T(x)$ and $(X-x)^p$; by 12.51

$$\frac{T(X)-T(x)}{-(X-x)^p} = \frac{T'(c)}{-p(X-c)^{p-1}}, \quad x < c < X,$$

and so

$$f(X) - T(x) = T(X) - T(x) = \frac{(X-x)^p}{p(X-c)^{p-1}} \frac{(X-c)^{n-1}}{(n-1)!} f^n(c),$$

whence $$f(x)-T(x) = h^n \frac{(1-\theta)^{n-p}}{p(n-1)!} f^n(x+\theta h),$$

which is the *Schlömilch* remainder. Cauchy's remainder is just a particular case of Schlömilch's with $p = 1$; if we take $p = n$ we obtain the *Lagrange* remainder

$$\frac{h^n}{n!} f^n(x+\theta h).$$

13.312. Applying Theorem 12.5 to the function $T(x)$ and any function $g(x)$ with derivative of constant sign we have

$$\frac{T(X)-T(x)}{g(X)-g(x)} = \frac{T'(c)}{g'(c)}, \quad x < c < X.$$

whence we obtain the remainder

$$f(X)-T(x) = \frac{g(X)-g(x)}{g'(c)} \frac{(X-c)^n}{n!} f^n(c)$$

which contains all the previous remainders as special cases.

Replacing the Cauchy by the Lagrange remainder, 13.31 takes the simpler form

13.32. $$f(x+h) = f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\dots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(x)+ \\ +\frac{h^n}{n!}f^n(x+\theta h).$$

For completeness we rewrite 13.31 also with the Schlömilch remainder, giving

13.33. $$f(x+h) = f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\dots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(x)+ \\ +h^n\frac{(1-\theta)^{n-p}}{p(n-1)!}f^n(x+\theta h).$$

The existence of a value of θ in $[0, 1]$ satisfying 13.31 or 13.312 or 13.32 or 13.33 is known as *Taylor's* theorem; the actual value of θ depends upon x, h, and n (and, of course, upon the function $f(x)$), and in general is not the same in the three equations.

The special case of Taylor's theorem with $x = 0$ is known as *Maclaurin's* theorem. Taking $x = 0$ in 13.32 we have

13.331. $$f(x) = f(0)+hf'(0)+\frac{h^2}{2!}f''(0)+\dots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(0)+ \\ +\frac{h^n}{n!}f^n(\theta x)$$

and from 13.33,

13.332. $$f(x)=f(0)+hf'(0)+\frac{h^2}{2!}f''(0)+\ldots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(0)+ +h^n\frac{(1-\theta)^{n-p}}{p(n-1)!}f^n(\theta x).$$

Since $f^n(x)$ is necessarily continuous, and $|\theta h| < |h|$, we can determine q, depending upon p, so that $f^n(x+\theta h) = f^n(x)+0(p)$ provided $h = 0(q)$. Hence, from 13.32, we find

13.34. $$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\ldots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(x + \frac{h^n}{n!}\{f^n(x)+0(p)\}$$

for any h satisfying $h = 0(q)$.

13.341. Conversely, if $f(x)$ is differentiable n times, and if for any h such that $h = 0(q)$, we have

13.342. $$f(x+h) = a_0(x)+ha_1(x)+\frac{h^2}{2!}a_2(x)+\ldots+\frac{h^n}{n!}\{a_n(x)+0(p)\}$$

then $a_0(x) = f(x)$ and for all r from 1 to n, $a_r(x) = f^r(x)$.

Proof. From 13.34 and 13.342 we have, for any h such that $h = 0(q)$,

$$\{a_0(x)-f(x)\}+h\{a_1(x)-f'(x)\}+\frac{h^2}{2!}\{a_2(x)-f''(x)\}+\ldots+ +\frac{h^n}{n!}\{a_n(x)-f^n(x)\} = \frac{h^n}{n!}0(p-1).$$

Take $h = 0$, then $a_0(x) = f(x)$. Suppose that $a_r(x) = f^r(x)$ for $r \leqslant k$, then

$$\frac{h^{k+1}}{(k+1)!}\{a_{k+1}(x)-f^{k+1}(x)\}+\ldots+\frac{h^n}{n!}\{a_n(x)-f^n(x)\} = \frac{h^n}{n!}0(p-1)$$

and therefore, provided $h \neq 0$, we may divide by h^{k+1}, giving, for $h = 0(q)$,

$$\frac{1}{(k+1)!}\{a_{k+1}(x)-f^{k+1}(x)\}+\ldots+\frac{h^{n-k-1}}{n!}\{a_n(x)-f^n(x)\} = \frac{h^{n-k-1}}{n!}0(p-1).$$

Denote the left-hand side of this equation by $\phi(h)$, so that $\phi(h)$ is

continuous in any interval and $\phi(h) = \dfrac{h^{n-k-1}}{n!}0(p-1)$ provided $h = 0(q)$, $h \neq 0$. Hence $\phi\left(\dfrac{1}{m}\right) = \left(\dfrac{1}{m}\right)^{n-k-1}\dfrac{1}{n!}0(p-1)$, and so $\left|\phi\left(\dfrac{1}{m}\right)\right| < \dfrac{1}{m}$ provided $k+1 < n$, proving $\phi\left(\dfrac{1}{m}\right) \to 0$; but as $\phi(h)$ is continuous, $\phi\left(\dfrac{1}{m}\right) \to \phi(0)$, and therefore $\phi(0) = 0$, i.e.

$$a_{k+1}(x) = f^{k+1}(x).$$

Thus we have proved that $a_r(x) = f^r(x)$ for all r less than n; but this implies that

$$a_n(x) - f^n(x) = 0(p-1)$$

for any p, and therefore $a_n(x) = f^n(x)$, which completes the proof.

13.343. Theorem 13.34 can be proved without an appeal to either Taylor's theorem or the mean-value theorem; it is in fact an immediate consequence of the following important result:

13.35. If $\phi(h)$ is differentiable n times and if

$$\phi(0) = \phi'(0) = \dots = \phi^{n-1}(0) = 0, \quad \text{and} \quad \phi^n(0) \neq 0,$$

then, for small values of $|h|$, $\phi(h)/h^n$ has the same sign as $\phi^n(0)$.

Suppose that $\phi^n(0) = 2\alpha > 0$. Since

$$\phi^{n-1}(h) - \phi^{n-1}(0) = h\{\phi^n(0) + 0(p)\}, \quad \text{provided } h = 0(q),$$

if we choose p so that $10^p\alpha > 1$, we find that $\phi^{n-1}(h)/h \geqslant \alpha$. Hence if $h > 0$, $\phi^{n-1}(h) \geqslant h\alpha$ and $-\phi^{n-1}(-h) \geqslant h\alpha$. Therefore $\int_0^h \{\phi^{n-1}(h) - h\alpha\}\, dh \geqslant 0$, i.e. $\phi^{n-2}(h) - \dfrac{h^2}{2!}\alpha \geqslant 0$; integrating a further $(n-2)$ times we arrive at the inequality $\phi(h) \geqslant \dfrac{h^n}{n!}\alpha$. Similarly, by integrating $-\phi^{n-1}(-h) - h\alpha$ repeatedly, we find

$$(-1)^n\phi(-h) \geqslant \frac{h^n}{n!}\alpha,$$

which may be written as $\phi(-h)/(-h)^n \geqslant \alpha/n!$. Thus whether h be positive or negative, provided $h = 0(q)$, $\phi(h)/h^n \geqslant \alpha/n!$.

In the same way we can show that if $\phi^n(0) = -2\beta < 0$, then $\phi(h)/h^n \leqslant -\beta/n!$, and 13.35 is proved.

13.351. Let λ be any constant and let

$$\phi(h) = f(x)+hf'(x)+\ldots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(x)+\frac{h^n}{n!}\{f^n(x)+\lambda\}-f(x+h).$$

Then $\phi^r(0) = 0$ for $r < n$, and $\phi^n(0) = \lambda$, and therefore, by 13.35, $\phi(h)/h^n$ has the same sign as λ, for $h = 0(q)$.

Replacing λ first by $1/10^k$ and then by $-1/10^k$ we see that, whether h^n be positive or negative,

$$f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\ldots+\frac{h^n}{n!}f^n(x)-f(x+h)$$

lies between $-\frac{h^n}{n!}\frac{1}{10^k}$ and $+\frac{h^n}{n!}\frac{1}{10^k}$, and therefore is equal to $\frac{h^n}{n!}0(k)$, $h = 0(q)$, which proves 13.34. Observe that the value of q for which the result holds depends upon the value of λ, and therefore upon k.

13.36. If h, k have the same signs, and if $f(x)$ is differentiable in an interval i which contains the points a, $a+h$, $a+k$ then there is a point c in i, such that

$$\mathbf{f(a+h)-f(a+k) = (h-k)f'(a)+(h^2-k^2)\frac{f''(a)}{2!}+}$$

$$\mathbf{+(h^3-k^3)\frac{f'''(a)}{3!}+\ldots+(h^{n-1}-k^{n-1})\frac{f^{n-1}(a)}{(n-1)!}+(h^n-k^n)\frac{f^n(c)}{n!}}.$$

Proof. Write

$$\phi(x) = f(x)-\left\{f(a)+(x-a)f'(a)+(x-a)^2\frac{f''(a)}{2!}+\ldots+(x-a)^{n-1}\frac{f^{n-1}(a)}{(n-1)!}\right\}$$

and $$\psi(x) = \frac{(x-a)^n}{n!};$$

then

$$\phi(a) = \phi'(a) = \phi''(a) = \ldots = \phi^{n-1}(a) = 0, \qquad \phi^n(x) = f^n(x),$$

and

$$\psi(a) = \psi'(a) = \psi''(a) = \ldots = \psi^{n-1}(a) = 0, \qquad \psi^n(x) = 1,$$

whence, by Theorem 12.52, we can find a point c in i such that

$$\frac{\phi(a+h)-\phi(a+k)}{\psi(a+h)-\psi(a+k)} = \frac{\phi^n(c)}{\psi^n(c)}$$

and so

$$f(a+h)-f(a+k) = (h-k)f'(a)+(h^2-k^2)\frac{f''(a)}{2!}+\dots+$$

$$+(h^{n-1}-k^{n-1})\frac{f^{n-1}(a)}{(n-1)!}+(h^n-k^n)\frac{f^n(c)}{n!}.$$

13.4. The endless Taylor series

In order that $f(X)$ be equal to its Taylor series it is both necessary and sufficient that the difference between $f(X)$ and $T_n(x)$, the sum to n terms of the series, be arbitrarily small for sufficiently great values of n. Expressing this difference by the Lagrange remainder, we require that there be a function N_p such that, for any p and $n \geqslant N_p$,

$$\frac{h^n}{n!}f^n(x+\theta h) = 0(p);$$

this condition, however, is of little practical value since θ varies with n in a manner that is unknown to us.

It is easy to deduce *sufficient* conditions which do not involve a knowledge of θ; for instance, if for any t between 0 and 1 and $n \geqslant N$

$$\frac{h^n}{n!}f^n(a+th) = 0(p)$$

(where N_p depends only upon p and *not* upon t), then, necessarily,

$$\frac{h^n}{n!}f^n(x+\theta h) = 0(p),$$

whatever the value of θ. But the condition is stronger than is necessary. *Pringsheim* has shown, however, that if we take the *Cauchy* remainder, instead of the Lagrange, the analogous condition, viz.

$$\frac{(1-t)^{n-1}}{(n-1)!}h^n f^n(a+th) = 0(p)$$

for any t in $[0, 1]$ and $n \geqslant N_p$, where N_p depends upon p but not upon t, is both *necessary* and *sufficient* for the equality of $f(x+h)$ and the Taylor series

$$f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+\dots+\frac{h^n}{n!}f^n(x)+\dots.\dagger$$

It is important to distinguish between the condition '$\phi_n(t) = 0(p)$,

† A proof of this important result is given in *The Theory of Functions of a Real Variable*, E. T. Hobson, vol. ii, p. 208.

The proof is simple and straightforward but rather too long to be given here.

for any t in $[0, 1]$ and $n \geqslant N(p)$' and the seemingly similar condition '$\phi_n(t) = 0(p)$, for any t in $[0, 1]$ and $n \geqslant N(p, t)$'.

For instance $(1-t)^n = 0(p)$ for any t in $[0, 1]$ and

$$n > p/\log_{10}\{1/(1-t)\},$$

but we cannot choose a value of n, independent of t, which for *any* t in $[0, 1]$, makes $(1-t)^n = 0(p)$, for $\left(1-\frac{1}{n}\right)^n \to 1/e$, and so however great n may be we can make $(1-t)^n$ near to $1/e$ by taking $t = 1/n$.

13.41. The convergence of the Taylor series of a function $f(X)$ does not ensure that the limit of the series is $f(X)$, for if the remainder after n terms tends, not to zero, but to some function $\phi(X)$, then the Taylor series is convergent, but its limit is

$$f(X)-\phi(X).$$

13.42. Simplest amongst the *sufficient* conditions for the equality of $f(x+h)$ and its Taylor series is the condition that for all n and all t in the open interval $[x, x+h]$

$$|f^n(t)| < M,$$

where M is a constant; for if this condition holds, then whatever value θ may have, in $[0, 1]$, the positive value of the Lagrange remainder $\frac{h^n}{n!}f^n(x+\theta h)$ is less than $M\frac{|h|^n}{n!}$, and $\frac{|h|^n}{n!} \to 0$ since the series $1+|h|+\frac{|h|^2}{2!}+\ldots$ is convergent (with limit $e^{|h|}$).

13.43. Generalization of the limit concept

If we can determine N_p, depending only upon p, so that, for any x in (a, b), $f_m(x)-f_n(x) = 0(p)$ provided $m, n \geqslant N_p$, then we say that the sequence $f_n(x)$ is *interval-* (or *uniformly*) *convergent in* (a, b).

13.44. If we can determine $f(x)$ and N_p such that for any x in (a, b), and $n \geqslant N_p$,

$$f_n(x) = f(x)+0(p)$$

then $f(x)$ is said to be the *interval-limit in* (a, b) of the sequence $f_n(x)$ and we write $f_n(x) \to f(x)$ in (a, b), or $\lim\limits_{(a,b)} f_n(x) = f(x)$ or, shortly, $\lim f_n(x) = f(x)$.

13.45. If $f_n(x)$ is interval-convergent in (a, b) then we can determine a function $f(x)$ such that $f(x)$ is the interval-limit in (a, b) of the sequence $f_n(x)$. For to any *fixed* x in (a, b) corresponds a definite convergent sequence $f_n(x)$ with a unique limit l; define $f(x)$ by the condition $f(x) = l$, so that for each x in (a, b), $f(x)$ is determined. It remains to show that $f(x)$ is the interval-limit in (a, b) of $f_n(x)$.

For a *chosen* x, $f_n(x)$ has the limit $f(x)$, i.e. for a chosen x,

$$f_n(x) = f(x)+0(p) \quad \text{for} \quad n \geqslant M,$$

where M depends not only upon p but also upon the chosen value of x, and so M may be written $M(p, x)$.

But for *any* x in (a, b), $f_m(x)-f_n(x) = 0(p)$, $m, n \geqslant N_p$, and so, taking $m \geqslant N_p$ and n not less than N_p or $M(p, x)$, we have

$$f_m(x) = f_n(x)+0(p) = f(x)+0(p)+0(p),$$

i.e.
$$f_m(x) = f(x)+0(p-1),$$

for *any* x in (a, b) and $m \geqslant N_p$, N_p depending only upon p, which proves that $f(x)$ is the interval-limit in (a, b) of $f_n(x)$.

13.46.* The proof of 13.45 brings to light an important distinction. Even if for any *chosen* x in (a, b), $f(x)$ is the limit of $f_n(x)$, it does not follow that $f(x)$ is the interval-limit in (a, b) of $f_n(x)$. For the condition that, for a chosen x, $f_n(x) \to f(x)$, involves only that $f_n(x) = f(x)+0(p)$ for $n \geqslant M$, where M depends upon the chosen value of x, as well as upon p, whereas to prove $\lim\limits_{(a,b)} f_n(x) = f(x)$ it is necessary that $f_n(x) = f(x)+0(p)$ for $n \geqslant N$, where N depends on p *alone* and not upon x.

Thus, for instance, if $f_n(x)$ is a convergent sequence for any *fixed* x in (a, b), the sequence $f_n(x)$, for this value of x, has a limit which depends upon x and is therefore a function, $f(x)$. But it is only for an assigned x that $f(x)$ is the limit of $f_n(x)$ and $f(x)$ is not necessarily the interval-limit in (a, b) of $f_n(x)$.

For example, if $f_n(x) = xn^2e^{-nx}$, then $f_n(0) = 0$ and so

$$\lim f_n(0) = 0,$$

and if x has some fixed value $c > 0$, then

$$f_n(c) = cn^2e^{-nc} < cn^2/(n^3c^3/3!) = 3!/c^2n,$$

and so $\lim f_n(c) = 0$. But it is not true that $\lim\limits_{(0,k)} f_n(x) = 0$ for

$$f_n\left(\frac{1}{n}\right) = \frac{1}{n}n^2e^{-1} = n/e,$$

and so however great n may be we can make $f_n(x) > 1$ by choosing $x = 1/n$, and therefore zero cannot be the interval-limit in $(0, k)$ of xn^2e^{-nx}.

13.47. If for any $n, f_n(x)$ is continuous in (a, b) and if $\lim\limits_{(a,b)} f_n(x) = f(x)$ then $f(x)$ is continuous in (a, b).

For, given any p, we can determine n_p so that $f(x) = f_{n_p}(x)+0(p)$, for any x in (a, b), and so $f(X)-f(x) = f_{n_p}(X)-f_{n_p}(x)+0(p-1)$. But $f_{n_p}(x)$ is continuous, and so $f_{n_p}(X)-f_{n_p}(x) = 0(p-1)$ provided $X-x = 0(q)$, whence $f(X)-f(x) = 0(p-2)$, for $X-x = 0(q)$, proving that $f(x)$ is continuous.

13.471. In the nomenclature of 13.43, Theorem 1.92 may be expressed by saying that if a power series $\sum a_n x^n$ is convergent at $x = X$, then $\sum a_n x^n$ is interval-convergent in the interval $(0, X)$.

13.5. If for any p, and any x, X in (a, b) such that $X-x = 0(q)$, $X \neq x$, where q depends only upon p,

13.501. $$f(X, x)-\phi(x) = 0(p),$$

then we say that $\phi(x)$ is the *interval-limit* in (a, b) of $f(X, x)$ as X *tends* to x, and write:

$$f(X, x) \to \phi(x) \quad \text{as} \quad X \to x \text{ in } (a, b), \quad \text{or} \quad \lim_{X\to x(a,b)} f(X, x) = \phi(x),$$

or shortly
$$\lim_{X\to x} f(X, x) = \phi(x).$$

13.51. If $f(X)$ does not depend upon x but only upon X and if, for any x, X in (a, b) such that $X-x = 0(q)$,

$$f(X)-\phi(x) = 0(p)$$

then $f(x)$ and $\phi(x)$ are *continuous* and *equal*; for if X_1, X_2 are any two points in (a, b) such that $X_1-X_2 = 0(q)$ and if $\bar{x}$ is the mid-point of X_1, X_2 then both $X_1-\bar{x} = 0(q)$, $X_2-\bar{x} = 0(q)$ and so

$$f(X_1)-\phi(\bar{x}) = 0(p), \qquad f(X_2)-\phi(\bar{x}) = 0(p),$$

whence $f(X_1)-f(X_2) = 0(p-1)$, which proves that $f(X)$ is continuous. Take any x, X such that $X-x = 0(q)$, then

$$f(X)-f(x) = 0(p-1) \quad \text{and} \quad f(X)-\phi(x) = 0(p)$$

so that, for any p, $f(x)-\phi(x) = 0(p-1)$, which proves that

$$\phi(x) = f(x).$$

13.52. If for some *definite* number l (as opposed to all the points of some interval)

$$f(X)-\phi(l) = 0(p)$$

for any $X \neq l$ such that $X-l = 0(q)$, q depending on p alone, then $\phi(l)$ is called the *point-limit* of $f(X)$ as X tends to l, and we write

$$\lim_{X\to l} f(X) = \phi(l);$$

but it does not follow that $f(l) = \phi(l)$, for the condition

$$f(X)-\phi(l) = 0(p)$$

imposes no restriction on the value of $f(X)$ at the point $X = l$.

It is important to observe in 13.5 that q depends only upon p and not upon x and X.

13.521. Theorem 13.34 may now be written in the form

$$f(x+h) = f(x)+hf'(x)+\frac{h^2}{2!}f''(x)+...+\frac{h^n}{n!}\{f^n(x)+\alpha_h\},$$

where $\lim_{h\to 0}\alpha_h = 0$.

For if

$$\alpha_h = \left[f(x+h)-\left\{f(x)+hf'(x)+...+\frac{h^n}{n!}f^n(x)\right\}\right]n!/h^n$$

then, by 13.34, $\alpha_h = 0(p)$ provided $h = 0(q)$, and so $\lim_{h\to 0}\alpha_h = 0$.

13.53. If for any p and any unequal x, X_1, X_2 in (a, b) such that

$$X_1-x = 0(q), \qquad X_2-x = 0(q),$$

where q depends only upon p,

$$f(X_1, x)-f(X_2, x) = 0(p)$$

then $f(X, x)$ is said to be *interval-convergent* in (a, b).

13.531. If $f(X, x)$ is interval-convergent in (a, b) then there is a function $\phi(x)$ such that $\lim_{X\to x(a,b)} f(X, x) = \phi(x)$.

Proof. We have, for any p, $f(X_1, x)-f(X_2, x) = 0(p)$, provided $X_1-x = 0(q)$, $X_2-x = 0(q)$, and so for any *fixed* x in (a, b),

$$f\left(x+\frac{1}{n}, x\right)-f\left(x+\frac{1}{m}, x\right) = 0(p), \quad \text{provided } n, m > 10^q,$$

which proves that the sequence $f\left(x+\frac{1}{n}, x\right)$ is convergent; let $\phi(x)$ be the limit of the sequence, and therefore for $n \geqslant N(p, x)$,

$$f\left(x+\frac{1}{n}, x\right)-\phi(x) = 0(p).$$

Take an X such that $X-x = 0(q)$ and an n greater than both 10^q and $N(p, x)$, then

$$f\left(x+\frac{1}{n}, x\right)-\phi(x) = 0(p)$$

and
$$f\left(x+\frac{1}{n}, x\right)-f(X, x) = 0(p),$$

whence

$$f(X, x)-\phi(x) = 0(p-1) \quad \text{for} \quad X-x = 0(q),$$

q depending only upon p, not on x, which proves

$$\lim_{X\to x(a,b)} f(X, x) = \phi(x).$$

13.54. If $f(X)$ is interval-convergent in (a, b) but does *not* contain x explicitly then $\lim\limits_{X\to x(a,b)} f(X) = f(x)$ and $f(x)$ is continuous.

For if $X_1-x = 0(q)$, $X_2-x = 0(q)$ then $f(X_1)-f(X_2) = 0(p)$, q depending only upon p, so that $f(x)$ is continuous and since $f(X)-f(x) = 0(p)$, provided only $X-x = 0(q)$, therefore

$$\lim_{X\to x(a,b)} f(X) = f(x).$$

13.55. If for some *definite* number l

$$f(X_1)-f(X_2) = 0(p)$$

for any X_1, $X_2 \neq l$ such that $X_1-l = 0(q)$, $X_2-l = 0(q)$, q depending only upon p, $f(x)$ is said to be point-convergent at l.

13.551. If $f(x)$ is point-convergent at l, then there is a number L such that $\lim\limits_{X\to l} f(X) = L$.

For $f(l+1/n)-f(l+1/m) = 0(p)$ provided $n, m > 10^q$, and so $f(l+1/n)$ is convergent; let L be its limit. Then

$$f\left(l+\frac{1}{n}\right)-L = 0(p) \quad \text{for} \quad n \geqslant N(p).$$

Hence for any X such that $X-l = 0(q)$ and for an n greater than both 10^q and $N(p)$ we have

$$f\left(l+\frac{1}{n}\right)-f(X) = 0(p),$$

and therefore $f(X)-L = 0(p)$ for any X such that $X-l = 0(q)$, which proves $\lim\limits_{X\to l} f(X) = L$.

EXAMPLES:

The condition that $f(x)$ be continuous in (a, b) is

(i) $$\lim_{X\to x(a,b)} f(X) = f(x),$$

and the relation between a differentiable function $f(x)$ and its derivative $f'(x)$ is

(ii) $$\lim_{X\to x(a,b)} \frac{f(X)-f(x)}{X-x} = f'(x).$$

(iii) If $\lim\limits_{X\to x(a,b)} P(X) = p(x)$ and $\lim\limits_{X\to x(a,b)} Q(X) = q(x)$ then

$$\lim_{X\to x(a,b)} \{P(X)+Q(X)\} = p(x)+q(x),$$

(iv) $$\lim_{X\to x(a,b)} P(X)Q(X) = p(x)q(x),$$

and if $q(x) \geqslant \lambda > 0$ in (a, b),

$$\lim_{X\to x(a,b)} P(X)/Q(X) = p(x)/q(x), \quad \text{etc.}$$

(the proofs are trivial and are left as an exercise to the reader).

13.56. If, for $X > 0$,

$$\lim_{X\to 0} f\left(\frac{1}{X}\right) = L, \quad \text{we write} \quad \lim_{X\to\infty} f(X) = L,$$

and if $$\lim_{X\to 0} f\left(-\frac{1}{X}\right) = L, \quad \text{we write} \quad \lim_{X\to-\infty} f(X) = L.$$

The sign '$\lim\limits_{X\to\infty} f(X)$' is read as 'the limit of $f(X)$ as X tends to infinity', and the sign '$\lim\limits_{X\to-\infty} f(X)$' as 'the limit of $f(X)$ as X tends to minus infinity'.

13.57. Extension of the interval concept

If $f(X)-f(x) = 0(p)$ provided $X-x = 0(q)$, where q depends only upon p, for any x, X such that $x \geqslant a$, $X \geqslant a$ we say that $f(x)$ is continuous in the interval $(a, \infty]$.

If $f(X)-f(x) = 0(p)$ provided $X-x = 0(q)$ for any x, X such that $x \leqslant b$, $X \leqslant b$ then $f(x)$ is said to be continuous in the interval $[-\infty, b)$, and if x, X may have any value whatsoever, provided only $X-x = 0(q)$, then we say that $f(x)$ is continuous in the interval $[-\infty, \infty]$.

Similarly, if

$$\{f(X)-f(x)\}/(X-x) = \phi(x)+0(p), \quad \text{provided } X-x = 0(q),$$

where q depends only upon p, for any x, X, such that $x \geqslant a$, $X \geqslant a$, then $f(x)$ is said to be differentiable in $(a, \infty]$, with analogous formulations for differentiability in $[-\infty, b)$ or in $[-\infty, \infty]$.

If $f_N(x)-f_n(x) = 0(p)$, $N > n \geqslant n_p$, and any $x \geqslant a$, where n_p depends only upon p, then $f_n(x)$ is said to be interval-convergent in $(a, \infty]$.

If the range of x is $x \leqslant b$ then $f_n(x)$ is interval-convergent in $[-\infty, b)$, and if x is totally unrestricted then $f_n(x)$ is said to be interval-convergent in $[-\infty, \infty]$.

Interval-convergence in $(0, X)$ for any X is not equivalent to interval-convergence in $(0, \infty]$; for instance, if $f_n(x) = x/(x+n)$, $N > n$ and $X > 0$, then in $(0, X)$

$$\begin{aligned} 0 < f_n(x)-f_N(x) &= x(N-n)/(x+n)(x+N) \\ &< x(N-n)/nN < x/n < X/n = 0(p), \end{aligned}$$

provided $n > 10^p X$, and so $f_n(x)$ is interval-convergent in $(0, X)$, for any X. But if x may have any positive value whatever, taking $x = n$, $N = 3n$, $f_n(x)-f_N(x) = 2n^2/8n^2 = \frac{1}{4}$, however great n may be, and so $f_n(x)$ is *not* interval-convergent in $(0, \infty]$.

13.6. L'Hospital's theorems

13.61. If $f(l) = g(l) = 0$ and for any $r < n$, $f^r(l) = g^r(l) = 0$ but $g^n(l) \neq 0$, then

$$\lim_{x \to l} \frac{f(x)}{g(x)} = \frac{f^n(l)}{g^n(l)}.$$

Write $x-l = h$; then by Taylor's theorem we can determine θ_1 and θ_2 in $[0, 1]$ such that

$$f(x) = f(l+h) = \frac{h^n}{n!} f^n(l+\theta_1 h)$$

and

$$g(x) = g(l+h) = \frac{h^n}{n!} g^n(l+\theta_2 h).$$

Since a derivative is continuous

$$\lim_{h\to 0} f^n(l+\theta_1 h) = f^n(l) \quad \text{and} \quad \lim_{h\to 0} g^n(l+\theta_2 h) = g^n(l),$$

and hence since $g^n(l) \neq 0$,

$$\lim_{x\to l}\frac{f(x)}{g(x)} = \lim_{h\to 0}\frac{f(l+h)}{g(l+h)} = \lim_{h\to 0}\frac{f^n(l+\theta_1 h)}{g^n(l+\theta_2 h)} = \frac{f^n(l)}{g^n(l)}.$$

13.62. If $f(l) = g(l) = 0$ and $g'(x)$ does not vanish in some interval containing the point l, except perhaps at l, and if $f'(x)/g'(x)$ tends to a limit as x tends to l, then

$$\lim_{x\to l}\frac{f(x)}{g(x)} = \lim_{x\to l}\frac{f'(x)}{g'(x)}.$$

For by the Cauchy formula, 12.5, we can determine a point c between l and x such that

$$\frac{f(x)-f(l)}{g(x)-g(l)} = \frac{f'(c)}{g'(c)},$$

and so, since $f(l) = g(l) = 0$, we have

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)};$$

but c lies between x and l and so $c \to l$ as $x \to l$, which proves that

$$\lim_{x\to l}\frac{f(x)}{g(x)} = \lim_{c\to l}\frac{f'(c)}{g'(c)}.$$

Theorem 13.62 is not just a special case of 13.61, as it may appear on the surface, for 13.62 may be deduced from 13.61 *only* if $g'(l) \neq 0$, whereas the proof we have given holds even if $g'(l) = 0$, though it requires instead that $g'(x) \neq 0$ for all values of x near l. Another point of interest is that the use of the Cauchy formula is essential, a double application of the mean-value theorem being insufficient; for if c_1 and c_2 tend to l independently we cannot prove that $f'(c_1)/g'(c_2)$ tends to a limit even if $f'(c)/g'(c)$ tends to a limit when c tends to l.

13.63. If $\lim\limits_{x\to\infty}\dfrac{f'(x)}{g'(x)}$ exists and $\lim f(x) = \lim g(x) = 0$, then

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to\infty}\frac{f'(x)}{g'(x)}.$$

For $$\lim_{x\to 0}\frac{f'(1/x)}{g'(1/x)} = \lim_{x\to\infty}\frac{f'(x)}{g'(x)} = L, \text{ say},$$
and so, by 12.51, for $X < c < x \leqslant x_p$,
$$\frac{f(1/X)-f(1/x)}{g(1/X)-g(1/x)} = \frac{f'(1/c)}{g'(1/c)} = L+0(p);$$
let $X \to 0$, then
$$\frac{f(1/x)}{g(1/x)} = L+0(p-1), \quad \text{i.e.} \quad \lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to\infty}\frac{f'(x)}{g'(x)}.$$
In order that the application of 12.51 may be valid, we require that $g'(1/x) \neq 0$ when x is near zero, i.e. that $g'(x) \neq 0$ for all sufficiently great values of x.

EXAMPLES. (i) $\lim\limits_{x\to 0}\dfrac{\sin x}{x} = 1$, for
$$\lim_{x\to 0}\frac{\sin x}{x} = \lim_{x\to 0}\frac{\cos x}{1} = 1.$$
(ii) $$\lim_{x\to 0}\frac{1-\cos x}{x^2} = \lim_{x\to 0}\frac{\cos x}{2} = \frac{1}{2}.$$
By repeated application of 13.63 we find
(iii) $$\lim_{x\to\infty}\frac{x^n}{e^{nx}} = \lim_{x\to\infty}\frac{n!}{n^n e^{nx}} = 0,$$
since $e^{nx} > nx$, if x is positive.

The conditions 13.61–3 for the existence of the limit of $f(x)/g(x$ are sufficient, but are *not* necessary, e g. $\lim\limits_{x\to\infty}\dfrac{\sin x}{x} = 0$, sinc $|\sin x| \leqslant 1$, but we cannot apply 13.63 since $\lim\limits_{x\to\infty}\dfrac{\cos x}{1}$ does not exist $\cos x$ varying from -1 to 1 as x increases from $(2n-1)\pi$ to $2n\pi$

13.64. If $f(x)$ and $g(x)$ are positive and differentiable near a, an if, as $x \to a$, $f(x) \to \infty$ (i.e. $1/f(x) \to 0$), $g(x) \to \infty$, and $f'(x)/g'(x) \to l$ then
$$\lim_{x\to a} f(x)/g(x) = l$$
provided $g'(x)$ does not vanish arbitrarily near to a.

Proof. Let x lie between a and X, then by Cauchy's formula
$$\frac{f(x)-f(X)}{g(x)-g(X)} = \frac{f'(\xi)}{g'(\xi)}, \quad \text{for a } \xi \text{ between } x, X,$$
wherefore $$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}\frac{\{1-g(X)/g(x)\}}{\{1-f(X)/f(x)\}}.$$

Choose X so near a that

$$l-1/k < f'(\xi)/g'(\xi) < l+1/k, \quad k \geqslant 2,$$

and keeping X fixed take x so near a that

$$|f(X)/f(x)| < 1/k, \qquad |g(X)/g(x)| < 1/k.$$

Then
$$\frac{k-1}{k+1} < \frac{1-g(X)/g(x)}{1-f(X)/f(x)} < \frac{k+1}{k-1}$$

and therefore

$$\left(l-\frac{1}{k}\right)\frac{k-1}{k+1} < \frac{f(x)}{g(x)} < \left(l+\frac{1}{k}\right)\frac{k+1}{k-1};$$

since both $\left(l-\frac{1}{k}\right)\frac{k-1}{k+1}$ and $\left(l+\frac{1}{k}\right)\frac{k+1}{k-1}$ tend to l, therefore

$$\lim_{x\to a} f(x)/g(x)$$

exists, and equals l.

13.65. If $f(x)$ and $g(x)$ are differentiable for all sufficiently great values of x, and if, as $x \to \infty$,

$$f(x) \to \infty, \quad g(x) \to \infty, \quad \text{and} \quad f'(x)/g'(x) \to l,$$

then
$$\lim_{x\to\infty} f(x)/g(x) = l$$

provided $g'(x)$ does not vanish for arbitrarily great values of x.

Proof. Let $x > X$, then by Cauchy's formula

$$\frac{f(x)-f(X)}{g(x)-g(X)} = \frac{f'(\xi)}{g'(\xi)}, \quad X < \xi < x.$$

Choose X so great that $l-1/k < f'(\xi)/g'(\xi) < l+1/k$, and keeping X fixed choose x so great that $|g(X)/g(x)| < 1/k$, $|f(X)/f(x)| < 1/k$, whence

$$(l-1/k)\frac{k-1}{k+1} < \frac{f(x)}{g(x)} < (l+1/k)\frac{k+1}{k-1}$$

and therefore $\lim_{x\to\infty} f(x)/g(x)$ exists, and equals l.

13.7. If $\lim_{X\to x}\dfrac{f(X)-g(X)}{(X-x)^n} = 0$, then the functions $f(X)$ and $g(X)$ are said to have *contact of the n-th order* at the point x.

13.71. If $f(X)$ and $g(X)$ have contact of the nth order then they have contact of the rth order, for any r less than n.

For

$$\lim_{X\to x}\frac{f(X)-g(X)}{(X-x)^r} = \lim_{X\to x}(X-x)^{n-r}\left\{\frac{f(X)-g(X)}{(X-x)^n}\right\} = 0.$$

13.72. If $f(X)$ and $g(X)$ are differentiable n times, the necessary and sufficient condition for $f(X)$ and $g(X)$ to have contact of the nth order at the point x is that $f(x) = g(x)$ and $f^r(x) = g^r(x)$, $r \leqslant n$.

For if $f^r(x) = g^r(x)$, we have by Taylor's theorem

$$\frac{f(x+h)-g(x+h)}{h^n} = \frac{1}{n!}\{f^n(x+\theta h)-g^n(x+\theta h)\} \to \frac{f^n(x)-g^n(x)}{n!}$$

$$\text{as } h \to 0$$

$$= 0;$$

and so writing $X = x+h$, we have $\lim_{X\to x} \dfrac{f(X)-g(X)}{(X-x)^n} = 0$, which proves that the condition is sufficient.

Conversely, if $\dfrac{f(x+h)-g(x+h)}{h^n} \to 0$, it follows that to any p corresponds a q such that, for $h = 0(q)$,

$$\{f(x)-g(x)\}+h\{f'(x)-g'(x)\}+\frac{h^2}{2!}\{f''(x)-g''(x)\}+\ldots+$$

$$+\frac{h^n}{n!}\{f^n(x+\theta h)-g^n(x+\theta h)\} = h^n 0(p)$$

and so, since $f^n(x+\theta h)-g^n(x+\theta h) = f^n(x)-g^n(x)+0(p)$, we have

$$\sum_{r=0}^{n} \frac{h^r}{r!}\{f^r(x)-g^r(x)\} = h^n 0(p-1).$$

Exactly as in 13.341 we can now prove that the coefficient of each power of h on the left-hand side is zero, and so $f^r(x) = g^r(x)$, for all r from 0 to n, which proves that the condition is necessary.

13.73. The terminating Taylor series

$$f(x)+(X-x)f'(x)+\frac{(X-x)^2}{2!}f''(x)+\ldots+\frac{(X-x)^n}{n!}f^n(x)$$

is the only polynomial of the nth degree which has contact of the nth order with the differentiable function $f(X)$ at the point x.

For if $\phi(X) = a_0(x)+(X-x)a_1(x)+\ldots+(X-x)^n a_n(x)$ has contact of the nth order with $f(X)$ at the point x, then

$$\phi^r(x) = f^r(x) \quad \text{for all} \quad r \leqslant n;$$

but

$$\frac{\phi^r(X)}{r!} = a_r(x)+\binom{r+1}{r}(X-x)a_{r+1}(x)+$$
$$+\binom{r+2}{r}(X-x)^2a_{r+2}(x)+...+\binom{n}{r}(X-x)^{n-r}a_n(x)$$

and so $\dfrac{\phi^r(x)}{r!} = a_r(x)$, which proves that $a_r(x) = \dfrac{f^r(x)}{r!}$, and so $\phi(X)$ is the terminating Taylor series associated with $f(X)$.

13.8. Maxima and minima

By means of Theorem 13.35 we can give an alternative proof of the criteria for maximum and minimum values of a function which we gave in § 7.75.

If $f'(a) = f''(a) = ... = f^{n-1}(a) = 0$ and $f^n(a) \neq 0$, then if n is even, $f(a)$ is a maximum or minimum value of $f(x)$ according as $f^n(a)$ is negative or positive, and if n is odd $f(a)$ is neither a maximum nor minimum value of $f(x)$; in other words $f(a)$ is a maximum or minimum value of $f(x)$ if and only if the first non-vanishing derivative $f^n(a)$ is of even order.

Proof. Let $\phi(h) = f(a+h)-f(a)$, then

$$\phi(0) = \phi'(0) \qquad = \phi^{n-1}(0) = 0, \quad \text{but} \quad \phi^n(0) = f^n(a) \neq 0.$$

Hence by 13.35, for sufficiently small values of h, $\{f(a+h)-f(a)\}/h^n$ has the same sign as $f^n(a)$; suppose that $f^n(a)$ is positive. Then $f(a+h)-f(a)$ has the same sign as h^n, and h^n is positive for any h, if n is even, and changes sign with h if n is odd. Similarly, if $f^n(a)$ is negative, $f(a+h)-f(a)$ has the sign opposite to h^n and is therefore negative if n is even, and changes sign with h if n is odd. Thus if n is even, $f(a+h) > f(a)$ for all sufficiently small values of $|h|$, when $f^n(a)$ is positive, so that $f(a)$ is a minimum value of $f(x)$, and when $f^n(a)$ is negative, $f(a+h) < f(a)$ and so $f(a)$ is a maximum value. If n is odd $f(a+h)-f(a)$ changes sign with h and so $f(a)$ is neither a maximum nor a minimum value.

EXAMPLE. Examine the behaviour of the function

$$(x-a)^m(x-b)^n,$$

at the points a, b, where $a > b$. Denote $(x-a)^m(x-b)^n$ by $\phi(x)$. Then by Leibnitz's theorem

$$\phi^p(x) = (x-a)^m D^p(x-b)^n + \binom{p}{1} D(x-a)^m . D^{p-1}(x-b)^n + \ldots +$$
$$+ \binom{p}{r} D^r(x-a)^m . D^{p-r}(x-b)^n + \ldots + (x-b)^n D^p(x-a)^m.$$

Since $D^r(x-a)^m = m(m-1)\ldots(m-r+1)(x-a)^{m-r}$, it follows that if $r < m$, $D^r(x-a)^m$ vanishes when $x = a$, and similarly if $s < n$, $D^s(x-b)^n$ vanishes when $x = b$; thus if $p \leqslant m$ we may write

$$\phi^p(x) = (x-b)^n D^p(x-a)^m + (x-a)\lambda(x),$$

and if $p \leqslant n$,

$$\phi^p(x) = (x-a)^m D^p(x-b)^n + (x-b)\mu(x),$$

where $\lambda(x)$ and $\mu(x)$ are polynomials in x.

In particular

$$\phi^n(b) = n!(b-a)^m \text{ and } \phi^p(b) = 0 \quad \text{if} \quad p < n,$$
$$\phi^m(a) = m!(a-b)^n \text{ and } \phi^p(a) = 0 \quad \text{if} \quad p < m.$$

Thus if m is odd the point a is a point of inflexion, and if n is odd the point b is a point of inflexion. If m is even then $\phi(a)$ is a minimum value, and if n is even, since $b-a$ is negative, $\phi(b)$ is a maximum if m is odd and a minimum if m is even.

13.9. Approximation to the root of an equation. Newton's method

If a is a root of the equation $f(x) = 0$ and if $\alpha_1, \alpha_2, \alpha_3, \ldots$ is a convergent sequence with limit a, then the decimals $\alpha_1, \alpha_2, \alpha_3, \ldots$ are called *successive approximations* to the root a. If $f(x)$ is a continuous function, and $\alpha_1, \alpha_2, \alpha_3, \ldots$ are successive approximations to a root a of the equation $f(x) = 0$, then the sequence $f(\alpha_1), f(\alpha_2), f(\alpha_3), \ldots$ is convergent and its limit is $f(a) = 0$; in other words, if $\alpha_1, \alpha_2, \alpha_3, \ldots$ are successive approximations to a root, $f(\alpha_1), f(\alpha_2), f(\alpha_3), \ldots$ tend to zero.

Any term in a sequence of successive approximations is called an approximation. If α is an approximation to a root a of an equation $f(x) = 0$ and if α^* is closer to a than α, i.e. if

$$|a-\alpha^*| < |a-\alpha|.$$

then α^* is said to be a *closer approximation* to the root than α. The difference $|a-\alpha|$ is called the *error* in the approximation α. The object of the present section is the determination of a sequence of successive approximations to a root of an equation $f(x) = 0$, when $f(x)$ is a differentiable function.

13.91. If x is a root of an equation $f(x) = 0$, and X an approximation to the root, then since $\dfrac{f(X)-f(x)}{X-x}$ is nearly equal to $f'(X)$ it follows that $X-x$ is nearly equal to $f(X)/f'(X)$, provided $f'(X) \neq 0$, and so x is nearly equal to $X-f(X)/f'(X)$; this suggests that $X-f(X)/f'(X)$ is a closer approximation to the root x than is X, but to show that this is actually the case we must be able to compare the errors in the two approximations, and this comparison is effected by means of Taylor's theorem.

Changing the notation, let a be an approximation, and $a+h$ the actual root, of an equation $f(x) = 0$, so that $|h|$ is the error in the approximation a. By Taylor's theorem

$$0 = f(a+h) = f(a)+hf'(a)+\frac{h^2}{2!}f''(a+\theta h).$$

We suppose that $f'(a) \neq 0$. Then the error in the approximation $a-f(a)/f'(a)$, viz. the positive value of $a+h-\{a-f(a)/f'(a)\}$, is equal to $\dfrac{h^2}{2!}|f''(a+\theta h)/f'(a)|$. If $f''(a+\theta h) = 0$, the error is zero; if not it is less than $|h|$ provided $|h| < 2|f'(a)/f''(a+\theta h)|$. Since $f'(a) \neq 0$, and $f''(x)$ being continuous, is bounded, this condition is possible. Hence if a is a sufficiently close approximation to the root of an equation $f(x) = 0$, then $a-f(a)/f'(a)$ is a closer approximation. This theorem is due to Newton.

13.92. It does *not* however follow, without further consideration, that, if $a_1 = a-f(a)/f'(a)$ and, for any n, $a_{n+1} = a_n-f(a_n)/f'(a_n)$, the sequence $a, a_1, a_2, a_3,\ldots$ is a sequence of successive approximations to the root $a+h$. For though Newton's theorem assures us that each a_{n+1} is a closer approximation to the root than is a_n, provided that the error in a_n is less than $2|f'(a_n)/f''(a_n+\theta_n h)|$, we have (as yet) no means of telling in advance whether this condition is satisfied for all values of n or not. It is, however, true that with

a sufficiently close initial approximation, the repeated application of Newton's process determines successive approximations to the root of an equation, but before we prove this we shall examine the general question of approximating to the root of an equation by a repetitive process.

13.93. *The sequence $c, f(c), f(f(c)), f\{f(f(c))\},\ldots$ determines a root of the equation $f(x) = x$, provided the sequence converges.*

For if $c_0 = c$ and $c_{n+1} = f(c_n)$, for all n, so that $c_1 = f(c)$, $c_2 = f(f(c))$, and so on, and if c_n converges to a limit σ, then

$$f(\sigma) = \lim f(c_n) = \lim c_{n+1} = \sigma$$

and σ is a root of the equation $f(x) = x$.

We shall consider a variety of conditions upon the function $f(x)$ which ensure the convergence of the sequence c_n.

13.931. If there is an interval (α, β), containing the point c, and such that $f(x)$ lies in (α, β) when x lies in (α, β), and if

$$|f'(x)| \leqslant M < 1$$

at all points of (α, β), then the sequence

$$c, \quad f(c), \quad f(f(c)), \quad \ldots$$

converges.

For by the mean-value theorem

$$|c_{n+1}-c_n| = |f(c_n)-f(c_{n-1})| = |c_n-c_{n-1}||f'(c_n^*)|,$$

where c_n^* lies in (c_n, c_{n-1}),

$$< M|c_n-c_{n-1}|,$$

whence $$|c_{n+1}-c_n| < M^n|c_1-c_0|;$$

but $\sum M^n$ converges, and so $\sum (c_{n+1}-c_n)$ is absolutely convergent, so that $c_0+\sum\limits_{r=0}^{n-1}(c_{r+1}-c_r) = c_n$ tends to a limit.

13.932. Writing $\phi(x) = x-f(x)$, so that $c_{n+1} = f(c_n) = c_n-\phi(c_n)$, then, if $\phi'(c)$ lies between $\frac{3}{4}$ and $\frac{5}{4}$ and if

$$|\phi''(x)| < 1/8|\phi(c)|$$

in the interval i_c with end-points $c\pm 2\phi(c)$,† the sequence c_n converges to a root of the equation $\phi(x) = 0$.

† More general conditions are: $\phi'(c)$ lies between $1\pm\delta$, and

$$|\phi''(x)| < (k-\delta)(1-k)/|\phi(c)|,$$

for all x between $c\pm\phi(c)/(1-k)$, where $0 < \delta < k < 1$.

Proof. We observe first that if x lies in i_c, then $|1-\phi'(x)| < \frac{1}{2}$, for

$$\phi'(x) = 1-\phi'(c)+\int \phi''(t)\,dt$$

and so $\quad |1-\phi'(x)| \leqslant \frac{1}{4}+|x-c|/8|\phi(c)| < \frac{1}{4}+\frac{1}{4} = \frac{1}{2}.$

Next we prove that for all n, c_n lies in i_c and $|\phi(c_n)| \leqslant |\phi(c)|/2^n$. This is true for $n = 0$, and if it is true for $n = 0, 1, 2,..., p$ then

$$|c_{p+1}-c| = \left|\sum_{r=0}^{p}(c_{r+1}-c_r)\right| \leqslant \sum_{r=0}^{p}|\phi(c_r)| \leqslant |\phi(c)|\sum_{r=0}^{p}1/2^r < 2|\phi(c)|,$$

so that c_{p+1} lies in i_c, and, furthermore,

$$\phi(c_{p+1})-\phi(c_p) = \int_{c_p}^{c_{p+1}} \phi'(t)\,dt$$

so that

$$\phi(c_{p+1}) = \int_{c_p}^{c_{p+1}} \{\phi'(t)-1\}\,dt, \quad \text{since} \quad \phi(c_p) = c_p-c_{p+1},$$

and therefore

$$|\phi(c_{p+1})| < \tfrac{1}{2}|c_{p+1}-c_p| = \tfrac{1}{2}|\phi(c_p)| \leqslant |\phi(c)|/2^{p+1}, \quad \text{by hypothesis,}$$

which proves our theorem for $n = p+1$, and so, by induction, it is true for all n. It follows that

$$|c_N-c_n| \quad \sum_{r=n}^{N-1}(c_{r+1}-c_r)\Big| \leqslant \sum_{r=n}^{N-1}|\phi(c_r)| \leqslant |\phi(c)|\sum_{r=n}^{N-1}1/2^r$$
$$< |\phi(c)|/2^{n-1},$$

and so the sequence c_n converges.

If σ is the limit of c_n, then

$$|\phi(\sigma)| = \lim|\phi(c_n)| \leqslant \lim|\phi(c)|/2^n = 0,$$

so that σ is a root of the equation $\phi(x) = 0$.

The error in the approximation c_n is

$$|\sigma-c_n \quad \lim_{N-} |c_N-c_n| \leqslant |\phi(c)|/2^{n-1}.$$

We can readily obtain other bounds for the error.

Since

$$c_{n+1}-c_n = -\phi(c_n) = \phi(\sigma)-\phi(c_n) = (\sigma-c_n)\phi'(\xi_n),$$

ξ_n lying between c_n and σ, therefore

$$|\sigma-c_n| = |c_{n+1}-c_n|/|\phi'(\xi_n)| < 2|c_{n+1}-c_n|,$$

showing that the error in c_n is less than $2|c_{n+1}-c_n|$.

It follows that

$$|\sigma - c_n| = |\sigma - c_{n-1} + c_{n-1} - c_n| < 3|c_n - c_{n-1}|$$

and so the error in c_n is less than $3|c_n - c_{n-1}|$.

13.933. In the foregoing theorem, if we take $\phi(x) = Ag(x)$, where A is constant, we see that the sequence c_n, given by

$$c_0 = c, \qquad c_{n+1} = c_n - Ag(c_n),$$

converges to a root of the equation $g(x) = 0$, provided that

$$\tfrac{3}{4} < Ag'(c) < \tfrac{5}{4} \quad \text{and} \quad |g''(x)| < 1/8A^2|g(c)|,$$

throughout the interval with end-points $c \pm 2Ag(c)$.

The arbitrary constant A enables us to satisfy the condition $\frac{3}{4} < Ag'(c) < \frac{5}{4}$ for any function $g(x)$, so long as $g'(c) \neq 0$. In particular we may take $A = 1/g'(c)$.

The successive approximation formula,

$$c_0 = c, \qquad c_{n+1} = c_n - Ag(c_n),$$

where A is near $1/g'(c)$, was first introduced by E. H. Neville, as a simplification of Newton's formula.

13.934. Taking $\phi(x) = f(x)/f'(x)$ we obtain Newton's original successive approximation formula

$$c_0 = c, \qquad c_{n+1} = c_n - g(c_n)/g'(c_n).$$

From 13.932 sufficient conditions for the convergence of c_n are seen to be

$$\frac{|f(c)f''(c)|}{\{f'(c)\}^2} < \frac{1}{4}, \qquad \left|\frac{d^2}{dx^2}\left(\frac{f(x)}{f'(x)}\right)\right| < \frac{|f'(c)|}{8|f(c)|}$$

in the interval bounded by the points $c \pm 2f(c)/f'(c)$.

These conditions are, however, unnecessarily complicated and a simpler convergence condition is given in the following theorem.

13.935. If $c_0 = c$, $c_{n+1} = c_n - f(c_n)/f'(c_n)$ and if

$$|f''(x)| < |f'(c)|^2/3|f(c)|$$

throughout the interval i_c bounded by the points $c \pm 2f(c)/f'(c)$ then c_n converges to a root of the equation $f(x) = 0$.

We prove first that all c_n lie in i_c and that

$$\left|\frac{f(c_n)}{f'(c_n)}\right| \leqslant \frac{1}{2^n}\left|\frac{2f(c)}{f'(c)}\right|.$$

For this is true when $n = 0$, and if it is true for $n = 1, 2,\dots, p$, then

$$|c_{n+1}-c| = \left|\sum_{r=0}^{p}(c_{r+1}-c_r)\right| \le \sum_{r=0}^{p}|f(c_r)/f'(c_r)|$$

$$\le \left|\frac{f(c)}{f'(c)}\right|\left(1+\frac{1}{2}+\frac{1}{2^3}+\frac{1}{2^7}+\dots\right) < 2\left|\frac{f(c)}{f'(c)}\right|$$

so that c_{p+1} lies in i_c.

By Taylor's theorem,

$$f(c_{p+1})-f(c_p)-(c_{p+1}-c_p)f'(c_p) = \tfrac{1}{2}(c_{p+1}-c_p)^2 f''(c_p^*),$$

c_p^* between c_p and c_{p+1},

and so

$$|f(c_{p+1})| = \tfrac{1}{2}(c_{p+1}-c_p)^2|f''(c_p^*)| \le \frac{1}{2}\cdot\frac{1}{3}\frac{|f'(c)|^2}{|f(c)|}\cdot 4\left|\frac{f(c)}{f'(c)}\right|^2\cdot\frac{1}{2^{2^{p+1}}}$$

$$= \frac{1}{2^{2^{p+1}}}\cdot\frac{2}{3}|f(c)|. \tag{i}$$

Furthermore, for any t in i_c,

$$|f'(t)-f'(c)| = \left|\int_c^t f''(x)\,dx\right| < |t-c|\cdot\tfrac{1}{3}|f'(c)|^2/|f(c)| < \tfrac{2}{3}|f'(c)|$$

and so $$\tfrac{1}{3}|f'(c)| < f'(t) < \tfrac{5}{3}f'(c). \tag{ii}$$

Hence, by (i) and (ii),

$$\left|\frac{f(c_{p+1})}{f'(c_{p+1})}\right| < \frac{1}{2^{2^{p+1}}}\,2\left|\frac{f(c)}{f'(c)}\right|$$

which proves the theorem for $n = p+1$, and so, by induction, for all values of n.

It follows that

$$|c_N-c_n| = \left|\sum_{r=n}^{N-1}(c_{r+1}-c_r)\right| \le \sum^{N-1}\left|\frac{f(c_r)}{f'(c_r)}\right|$$

$$< 2\left|\frac{f(c)}{f'(c)}\right|\left\{\frac{1}{2^{2^n}}+\frac{1}{2^{2^{n+1}}}+\dots\right\} < \frac{\ }{2^{2^n}}\left|\frac{f(c)}{f'(c)}\right|$$

showing that c_n is convergent. If σ is the limit of c_n then

$$|f(\sigma)| = \lim|f(c_n)| \le \lim\frac{1}{2^{2^n}}\left|\frac{2f(c)}{f'(c)}\right||f'(c_n)|$$

$$\le \lim\frac{1}{2^{2^n}}\left|\frac{2f(c)}{f'(c)}\right|\cdot\frac{5}{3}|f'(c)| = 0,$$

by (ii), so that σ is a root of the equation $f(x) = 0$.

The error in the approximation c_n is

$$|\sigma - c_n| = \lim_{N\to\infty} |c_N - c_n| \leqslant 2\left|\frac{f(c)}{f'(c)}\right|\left\{\frac{1}{2^{2^n}} + \frac{1}{2^{2^{n+1}}} + \dots\right\} < \frac{4}{2^{2^n}}\frac{f(c)}{f'(c)}$$

In terms of the difference of two successive approximations, a bound for the error in c_{n+1} is

$$\tfrac{3}{2}\{M/|f'(c)|\}(c_{n+1}-c_n)^2,$$

where M is a bound of $f''(x)$ in i_c (so that $M \leqslant \frac{1}{3}|f'(c)/f(c)|$).

For $\quad |f(c_{n+1})| = \frac{1}{2}(c_{n+1}-c_n)^2|f''(c_n^*)| < \frac{1}{2}(c_{n+1}-c_n)^2 M$

and

$$|f(c_{n+1})| = |f(c_{n+1})-f(\sigma)| = |(c_{n+1}-\sigma)f'(\xi_{n+1})| \geqslant \tfrac{1}{3}|c_{n+1}-\sigma||f'(c)|$$

so that $\qquad |c_{n+1}-\sigma| < \frac{3}{2}\{M/|f'(c)|\}(c_{n+1}-c_n)^2.$

Contrasting these expressions for the error in c_n with the corresponding expressions in § 13.932, it is evident that Newton's sequence converges more rapidly to the root than, for instance, the sequence of § 13.933, but in practical computation this advantage is sometimes outweighed by the simplicity of 13 933, where a *constant* factor A takes the place of the *changing* denominator $f'(c_n)$ in Newton's formula.

EXAMPLES. I. To approximate to the positive root of the equation $x^2-2 = 0$ by Newton's formula.

We have $f(x) = x^2-2$, $f'(x) = 2x$, $f''(x) = 2$.

Take $c = 1\frac{1}{2}$, then $f(c) = 1/4$, $f'(c) = 3$, and so $\frac{1}{3}\{f'(c)\}^2/f(c) = 12$, showing that the convergence condition $|f''(x)| < \frac{1}{3}\{f'(c)\}^2/|f(c)|$ is amply satisfied.

The successive approximations are given by

$$c = c_0 = 1\tfrac{1}{2}, \quad c_1 = 1\tfrac{1}{2}-\{(1\tfrac{1}{2})^2-2\}/2.1\tfrac{1}{2} = 17/12 = 1{\cdot}417\dots,$$

$$c_2 = 17/12-\{(17/12)^2-2\}/2(17/12) = 577/408 = 1{\cdot}414215\dots, \text{ etc.}$$

Since $\frac{3}{2}|f''(x)/f'(c)| = 1$, and $c_1-c_2 = 1/408$, therefore the error in the approximation c_2 is less than $1/(408)^2$, which is less than 10^{-5}, so that $\sqrt{2} = 1{\cdot}4142$ to four places.

II. To approximate to the root of the equation $x-\log x = 2$ by Neville's formula.

Take $c = 3$, $A = \frac{3}{2}$. Since $f(x) = x-\log x-2$, therefore

$$f'(x) = 1-1/x, \qquad f''(x) = 1/x^2,$$

and the interval i_c is (2·7062..., 3·2958...), which is contained between 2·7 and 3·3; furthermore $Af'(c) = 1$, $8A^2|f(c)| = 2{\cdot}3748$ and so, in i_c, $|f''(x)| < 1/(2{\cdot}7)^2 < 1/8A^2|f(c)|$, and the convergence conditions are amply satisfied.

The successive approximations are

$$c_1 = 3-3f(3)/2 = 3{\cdot}1479..., \qquad c_2 = 3{\cdot}1479-3f(c_1)/2 = 3{\cdot}1463...,$$
$$c_3 = 3{\cdot}1463-3f(c_2)/2 = 3{\cdot}1460....$$

The error in the approximation c_3 is less than $3|c_2-c_3|$, i.e. less than ·0009.

III. To approximate to the root of the equation $\cos x = x$.

In the interval (0, 1) $\cos x$ itself lies in (0, 1), and

$$\left|\frac{d}{dx}\cos x\right| \leq \sin 1 < 1$$

Hence the conditions of 13.931 are satisfied and

$$c, \quad c_1 = \cos c, \quad c_2 = \cos c_1, \quad c_3 = \cos c_2, \quad ...$$

are successive approximations to the root of the equation $\cos x = x$, for any c in (0, 1). Taking $c = \frac{1}{4}\pi = {\cdot}7854...$, we have in turn

$$c_1 = \cos\tfrac{1}{4}\pi = {\cdot}7071..., \qquad c_2 = \cos{\cdot}7071 = {\cdot}7603, \qquad c_3 = {\cdot}7246,$$
$$...., \qquad c_{14} = {\cdot}7393, \qquad c_{15} = {\cdot}7390.$$

XIV

CONSTRUCTIVE DEFINITION OF THE INTEGRAL

THE INTEGRAL AS LIMIT OF A CONVERGENT SEQUENCE. APPROXIMATION TO AN INTEGRAL. THE INTEGRAL AND DERIVATIVE OF A SEQUENCE OF FUNCTIONS

14. If $a = u_0, u_1, u_2, \ldots, u_{p+1} = b$ and $a = v_0, v_1, v_2, \ldots, v_{q+1} = b$ are two chains of points from a to b such that

$$|f(x)-f(X)| < M$$

for any two points x, X in the same sub-interval (u_r, u_{r+1}), and

$$|f(x)-f(X)| < N$$

for any two points x, X in the same sub-interval (v_s, v_{s+1}), and if

$$U = f(u_0)(u_1-u_0)+f(u_1)(u_2-u_1)+\ldots+f(u_r)(u_{r+1}-u_r)+\ldots+ \\ +f(u_p)(u_{p+1}-u_p)$$

and

$$V = f(v_0)(v_1-v_0)+f(v_1)(v_2-v_1)+\ldots+f(v_s)(v_{s+1}-v_s)+\ldots+ \\ +f(v_q)(v_{q+1}-v_q)$$

then
$$|U-V| < (b-a)(M+N).$$

Let $a = w_0, w_1, \ldots, w_{\mu+1} = b$ be the chain formed by all the points (u_r) and (v_s), and suppose that $w_{i(r)}$ and $w_{j(r)}$ are what the points u_r and u_{r+1} become in the new enumeration. Then, if

$$W = \sum_{t=0}^{\mu} f(w_t)(w_{t+1}-w_t),$$

we have

$$U-W = \sum_{r=0}^{p} f(u_r)(u_{r+1}-u_r) - \sum_{t=0}^{\mu} f(w_t)(w_{t+1}-w_t)$$

$$= \sum_{r=0}^{p} f(u_r)\Big[\sum_{s=i}^{j-1}(w_{s+1}-w_s)\Big] - \sum \Big\{\sum_{s=i}^{j-1} f(w_s)(w_{s+1}-w_s)\Big\}$$

$$= \sum \Big\{\sum_{s=i}^{j-1} f(w_i)(w_{s+1}-w_s)\Big\} - \sum \Big\{\sum_{s=i}^{j-1} f(w_s)(w_{s+1}-w_s)\Big\},$$

the outer sums extending over all the intervals

$$(u_r, u_{r+1}) = (w_i, w_j),$$

$$= \sum \Big[\sum_{s=i}^{j-1} \{f(w_i)-f(w_s)\}(w_{s+1}-w_s)\Big].$$

But w_s, for $i(r) \leqslant s \leqslant j(r)$, are all points of the interval (u_r, u_{r+1}), so that

$$|f(w_i)-f(w_s)| < M$$

and therefore

$$|U-W| < \sum \left\{M \sum_{s=i}^{j-1} (w_{s+1}-w_s)\right\}$$

$$= \sum_{r=0}^{p} \{M(u_{r+1}-u_r)\} = M(b-a).$$

Similarly, $|V-W| < N(b-a)$, and therefore

$$|U-V|$$

$$= |(U-W)+(W-V)| \leqslant |U-W|+|V-W| < (b-a)(M+N).$$

14.1. The function $f(x)$ is continuous in (a, b).

If $$a = u_0^k,\ u_1^k,\ u_2^k,\dots,\ u_{p_k+1}^k = b$$

is a chain of points from a to b such that

$$f(X)-f(x) = 0(k)$$

for any two x, X in the same sub-interval (u_r^k, u_{r+1}^k), and if

$$S_k = f(u_0^k)(u_1^k-u_0^k)+f(u_1^k)(u_2^k-u_1^k)+\ ..+f(u_{p_k}^k)(u_{p_k+1}^k-u_{p_k}^k)$$

then the sequence $S_1, S_2, S_3,\dots$ is convergent.

For by Theorem 14, if $m > n$,

$$S_m-S_n = (b-a)[0(m)+0(n)] = (b-a)0(n-1),$$

which proves that the sequence $S_1, S_2, S_3,\dots$ is convergent.

14.11. A chain of points $a = a_0, a_1, a_2,\dots, a_{p+1} = b$, such that $f(X)-f(x) = 0(n)$ for any two x, X in the same sub-interval (a_r, a_{r+1}), is called an 'n-chain' of the function $f(x)$.

14.12. If both $a = u_0^n,\ u_1^n,\ u_2^n,\dots,\ u_{p_n+1}^n = b$ and $a = v_0^n,\ v_1^n,\dots,$ $v_{q_n+1}^n = b$ are n-chains of the function $f(x)$, and if

$$S_n = \sum f(u_r^n)(u_{r+1}^n-u_r^n), \qquad T_n = \sum f(v_r^n)(v_{r+1}^n-v_r^n),$$

then the sequences $S_1, S_2, S_3,\dots$ and $T_1, T_2, T_3,\dots$ are both convergent and have the same limit.

The convergence of the sequences is given by 14.1; let S and T be the limits of (S_n) and (T_n).

By Theorem 14,

$$S_n-T_n = (b-a)[0(n)+0(n)] = (b-a)0(n-1).$$

But, if $m > n$,

$$S_m - S_n = (b-a)0(n-1) \quad \text{and} \quad T_m - T_n = (b-a)0(n-1),$$

and for a sufficiently great m,

$$S - S_m = (b-a)0(n-1) \quad \text{and} \quad T - T_m = (b-a)0(n-1),$$

so that

$$S - S_n = (b-a)0(n-2), \qquad T - T_n = (b-a)0(n-2)$$

and therefore

$$\begin{aligned} S-T &= (S-S_n)-(T-T_n)+S_n-T_n \\ &= (b-a)\{0(n-2)+0(n-2)+0(n-1)\} = (b-a)0(n-3), \end{aligned}$$

and this is true for any n, so that $S = T$.

14.2. If $a = a_0, a_1, a_2,..., a_{p+1} = b$ is an n-chain of a continuous function $f(x)$ and $S_n = \sum_{r=0}^{p} f(a_r)(a_{r+1}-a_r)$, and if $S_n \to S$, then S is called the *definite integral of* $f(x)$ *from* a *to* b, and $-S$ the definite integral from b to a.

By 14.1, the sequence (S_n) is convergent, and so the limit S exists, and by 14.12, if the sum T_n is formed on any other n-chain, then $T_n \to S$, and so the limit S is *unique*, and *independent* of the particular sequence of chains by which it is determined.

In the following sections we prove the equivalence of this definition of a definite integral with that given in Chapter IX. Anticipating the proof of this equivalence we denote the definite integral, as defined in 14.2, by the sign $\int_a^b f(x)\,dx$.

14.21. If $f(x)$ has the constant value k in (a, b) then

$$\int_a^b k\,dx = k(b-a).$$

For $f(x)$ is continuous, and if $a = a_0, a_1, a_2,..., a_{p+1} = b$ is an n-chain of $f(x)$, then

$$\begin{aligned} S_n &= f(a_0)(a_1-a_0)+f(a_1)(a_2-a_1)+...+f(a_p)(a_{p+1}-a_p) \\ &= k\{(a_1-a_0)+(a_2-a_1)+...+(a_{p+1}-a_p)\} = k(b-a) \end{aligned}$$

and so $S_n \to k(b-a)$, proving 14.21.

14.22. If $f(x)$ is continuous in (a, b), $a = a_0, a_1,\ldots, a_{p+1} = b$ is an n-chain of $f(x)$, and c_r is *any* point in (a_r, a_{r+1}), and if

$$S_n^* = f(c_0)(a_1-a_0)+f(c_1)(a_2-a_1)+f(c_2)(a_3-a_2)+\ldots+f(c_p)(a_{p+1}-a_p)$$

then
$$S_n^* \to \int_a^b f(x)\,dx.$$

For if
$$S_n = f(a_0)(a_1-a_0)+\ldots+f(a_p)(a_{p+1}-a_p)$$

then
$$S_n - S_n^* = \sum_{r=0}^{p} \{f(a_r)-f(c_r)\}(a_{r+1}-a_r)$$
$$= 0(n) \sum_{r=0}^{n} (a_{r+1}-a_r) = 0(n)(b-a),$$

since both c_r and a_r lie in (a_r, a_{r+1}), and (a_r) is an n-chain.

But
$$\int_a^b f(x)\,dx - S_n = 0(n-2)(b-a)$$

and so
$$\int_a^b f(x)\,dx - S_n^* = 0(n-3)(b-a),$$

which proves that $S_n^* \to \int_a^b f(x)\,dx.$

14.23. If $f(x)$ is continuous in (a, b) and if c lies in (a, b), then

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx.$$

Let $a = a_0, a_1,\ldots, a_{p+1} = c$ and $c = a_{p+1}, a_{p+2},\ldots, a_{p+q+1} = b$ be n-chains of $f(x)$ so that $a = a_0, a_1,\ldots, a_{p+q+1} = b$ is an n-chain, and let

$$S_n = \sum_{r=0}^{p+q} f(a_r)(a_{r+1}-a_r),$$

$$S_n^1 = \sum_{r=0}^{p} f(a_r)(a_{r+1}-a_r), \qquad S_n^2 = \sum_{r=p+1}^{p+q} f(a_r)(a_{r+1}-a_r),$$

so that
$$S_n \to \int_a^b f(x)\,dx, \qquad S_n^1 \to \int_a^c f(x)\,dx, \quad \text{and} \quad S_n^2 \to \int_c^b f(x)\,dx;$$

but $S_n = S_n^1+S_n^2$ and therefore $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$

14.24. If $f(x)$ and $g(x)$ are continuous in (a, b) then

$$\int_a^b \{f(x) \pm g(x)\}\, dx = \int_a^b f(x)\, dx \pm \int_a^b g(x)\, dx.$$

Since $f(x)$ and $g(x)$ are continuous, therefore $f(x) \pm g(x)$ is continuous and so we can determine a chain $a = a_0, a_1, a_2, \ldots, a_{p+1} = b$ which is an n-chain for all four functions $f(x)$, $g(x)$, $f(x) \pm g(x)$; for instance, by combining an $(n+1)$-chain of $f(x)$ with an $(n+1)$-chain of $g(x)$ we obtain an n-chain of the four functions.

Hence if

$$S_n^1 = \sum_0^p f(a_r)(a_{r+1} - a_r), \qquad S_n^2 = \sum_0^p g(a_r)(a_{r+1} - a_r)$$

and

$$S_n = \sum^p \{f(a_r) \pm g(a_r)\}(a_{r+1} - a_r),$$

then

$$S_n^1 \to \int_a^b f(x)\, dx, \qquad S_n^2 \to \int_a^b g(x)\, dx, \quad \text{and} \quad S_n \to \int_a^b \{f(x) \pm g(x)\}\, dx;$$

but $S_n = S_n^1 \pm S_n^2$, whence 14.24 follows.

14.25. If $f(x)$ is continuous in (a, b) and $f(x) \geqslant 0$, then

$$\int^b f(x)\, dx \geqslant 0.$$

For

$$S_n = f(a_0)(a_1 - a_0) + f(a_1)(a_2 - a_1) + \ldots + f(a_p)(a_{p+1} - a_p) \geqslant 0$$

and so $\lim S_n \geqslant 0$.

14.251. If $f(x)$ is continuous, so that $f(x)$ is bounded and lies between m and M say, then

$$m \leqslant \frac{1}{b-a} \int_a^b f(x)\, dx \leqslant M.$$

For $M - f(x) \geqslant 0$, so that by 14.25

$$\int^b \{M - f(x)\}\, dx \geqslant 0,$$

and hence, by 14.24,

$$\int^b f(x)\, dx \leqslant \int^b M\, dx = M(b-a).$$

Similarly, $$\int_a^b f(x)\,dx \geqslant m(b-a).$$

14.252. If $f(x)$ is continuous in (a, b), $a < b$, then
$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx.$$

Since $|f(x)-f(X)| \geqslant ||f(x)|-|f(X)||$, therefore $|f(x)|$ is continuous and therefore integrable. Moreover
$$\left|\sum f(a_r)(a_{r+1}-a_r)\right| \leqslant \sum |f(a_r)|(a_{r+1}-a_r),$$
whence
$$\left|\int_a^b f(x)\,dx\right| = \lim \left|\sum f(a_r)(a_{r+1}-a_r)\right|$$
$$\leqslant \lim \sum_0^p |f(a_r)|(a_{r+1}-a_r) = \int_a^b |f(x)|\,dx.$$

14.3. If f(x) is continuous in (a, b), and if t lies in (a, b), then $\int_a^t$ f(x) dx is a differentiable function of t, in (a, b), and its derivative is f(t).

The value of the integral $\int_a^t f(x)\,dx$ is, by definition, dependent on the value of t, and may therefore be denoted by $F(t)$. Observe that for any t, T
$$F(T)-F(t) = \int_a^T f(x)\,dx - \int_a^t f(x)\,dx = \int_t^T f(x)\,dx.$$

Since $f(x)$ is continuous in (a, b) we can determine k_n so that
$$f(T)-f(t) = 0(n)$$
for any t, T in (a, b) such that $T-t = 0(k_n)$; in particular if x lies between t and T, $f(x)-f(t) = 0(n)$.

Now $f(x)-f(t)$ is a continuous function of x, and so if
$$0 < T-t = 0(k_n),$$
$$\left|\left[\int_t^T \{f(x)-f(t)\}\,dx\right]\Big/(T-t)\right| \leqslant \left\{\int_t^T |f(x)-f(t)|\,dx\right\}\Big/(T-t),$$
by 14.252,
$$= 0(n), \quad \text{by 14.251},$$

and similarly, if $0 < t-T = 0(k_n)$,

$$\left|\left[\int_t^T \{f(x)-f(t)\}\,dx\right]\Big/(T-t)\right| = \left|\int_T^t \{f(x)-f(t)\}\,dx\Big/(t-T)\right|$$

$$\leqslant \int_T^t |f(x)-f(t)|\,dx\,/(t-T)$$

$$= 0(n);$$

but $$\int_t^T \{f(x)-f(t)\}\,dx = \int_t^T f(x)\,dx - (T-t)f(t),$$

and therefore $$\frac{1}{T-t}\int_t^T f(x)\,dx - f(t) = 0(n),$$

i.e. $$\frac{F(T)-F(t)}{T-t} - f(t) = 0(n), \quad \text{provided } T-t = 0(k_n),$$

which proves that $f(t)$ is the derivative of $F(t)$ in (a, b).

14.31. If $G(t)$ is any function whose derivative in (a, b) is $f(t)$ then

$$\int_a^t f(x)\,dx = G(t)-G(a).$$

Let $\int_a^t f(x)\,dx = F(t)$, so that, by 14.3, $F(t)$ and $G(t)$ have the same derivative, and therefore $G(t) = F(t)+C$, i.e.

$$G(t)-C = \int_a^t f(x)\,dx.$$

Since $|f(x)|$ is continuous in (a, b) it is bounded, by M say, and therefore

$$\left|\int_a^t f(x)\,dx\right| \leqslant |t-a|M = 0(n+1) \quad \text{if} \quad |t-a| < 1/M.10^{n+1}.$$

Furthermore, $G(t)$ is differentiable and therefore continuous, and so we can determine μ_n so that

$$G(t)-G(a) = 0(n+1) \quad \text{provided } |t-a| < 1/\mu_n.$$

Hence if λ_n is the greater of μ_n and $M.10^{n+1}$ and $|t-a| < 1/\lambda_n$ then

$$|G(a)-C| = |G(t)-C+G(a)-G(t)|$$
$$\leqslant |G(t)-C|+|G(a)-G(t)| = 0(n+1)+0(n+1) = 0(n).$$

Since this is true for any n, $G(a)-C = 0$ and so

$$\int_a^t f(x)\,dx = G(t)-G(a),$$

i.e.
$$\int_a^t G'(x)\,dx = G(t)-G(a).$$

This completes the proof of the equivalence of the definite integral as limit of a sum and the definite integral as the operation inverse to differentiation.

The difference between the two definitions is, however, as important as the equivalence theorem we have just established. For the integral which is but the inverse of a derivative is known to exist only after we have found it—and this requires a previous knowledge of the integral function—whereas the integral which is a limit of a sum is *constructible*, and therein lies the novelty and importance of the definition of this chapter. Previously we could only say that a derivative is integrable, in the sense that if one function is known to be the derivative of another then the latter is the integral of the first, whereas, by 14.2, we can now say that any continuous function is integrable, and we can construct its integral.

14.4. The origin of the definitions of area and length which we gave in §§ 10.34 and 10.41 is to be found in the constructive definition of an integral.

If $f(x)$ is continuous in (a, b) and if $a = a_0, a_1, a_2,..., a_{p+1} = b$ is an n-chain of $f(x)$, we define a function $\phi_n(x)$ by the condition:

for each r from 0 to p, $\phi_n(x) = f(a_r)$ for any x such that

$$a_r \leqslant x < a_{r+1}.$$

Thus since any x in (a, b) must lie in one of the sub-intervals (a_r, a_{r+1}), the function $\phi_n(x)$ is defined for any x in (a, b). $\phi_n(x)$ is called a *step-function.*

Since $\phi_n(x)$ is constant in each of the intervals $(a_r, a_{r+1}]$, the broken line $y = \phi_n(x)$, the x-axis, and the ordinates at each of the points $x = a_r$ determine a set of rectangles; the rectangle bounded by $y = \phi_n(x)$, the x-axis, and the ordinates of a_r and a_{r+1} is of height

$f(a_r)$ and breadth $a_{r+1}-a_r$, and so the area bounded by $y = \phi_n(x)$, the x-axis, and the ordinates at a and b is

$$S_n = f(a_0)(a_1-a_0)+f(a_1)(a_2-a_1)+\ldots+f(a_p)(a_{p+1}-a_p)$$

and $S_n \to \int_a^b f(x)\,dx$.

But $f(x)-\phi_n(x) = 0(n)$, for each x in (a, b) lies in some sub-interval (a_r, a_{r+1}) and so $f(x)-\phi_n(x) = f(x)-f(a_r)$, $a_r \leqslant x < a_{r+1}$. Hence $\phi_n(x) \to f(x)$. It is therefore natural to define the area bounded by the curve $y = f(x)$ to be the limit of the area bounded by $y = \phi_n(x)$, i.e. $\int_a^b f(x)\,dx$, which gives 10.34.

In a similar way we can establish the formula 10.41 for the length of an arc. Let $f(x)$ be a differentiable function and $a = a_0, a_1,\ldots, a_{p+1} = b$ an n-chain of the continuous function

$$\sqrt{[1+\{f'(x)\}^2]}.$$

The distance between the points $(a_r, f(a_r))$, $(a_{r+1}, f(a_{r+1}))$ is $\{(a_{r+1}-a_r)^2+(f(a_{r+1})-f(a_r))^2\}^{\frac{1}{2}}$, and so the length of the open polygon from $(a, f(a))$ to $(b, f(b))$ with vertices at the points $(a_r, f(a_r))$, $r = 1, 2, 3,\ldots, p$, is

$$L_n = \sum_{r=0}^{p} \{(a_{r+1}-a_r)^2+(f(a_{r+1})-f(a_r))^2\}^{\frac{1}{2}}.$$

By the mean-value theorem, we can determine c_r in (a_r, a_{r+1}) such that

$$f(a_{r+1})-f(a_r) = (a_{r+1}-a_r)f'(c_r),$$

and so if $\phi(x) = \sqrt{\{1+(f'(x))^2\}}$, the length of the polygon takes the form

$$L_n = \sum_{r=0}^{p} \phi(c_r)(a_{r+1}-a_r)$$

and
$$L_n \to \int_a^b \phi(x)\,dx = \int_a^b \sqrt{[1+\{f'(x)\}^2]}\,dx.$$

The equation of the line joining the points $(a_r, f(a_r))$, $(a_{r+1}, f(a_{r+1}))$ is

$$\{y-f(a_r)\}/(x-a_r) = \{f(a_{r+1})-f(a_r)\}/(a_{r+1}-a_r) = f'(c_r),$$

i.e. $y = f(a_r)+(x-a_r)f'(c_r)$, and so at any point between a_r and a_{r+1} the difference between the value of y on this line and on the curve $y = f(x)$ is $f(a_r)-f(x)+(x-a_r)f'(c_r)$. Let M be a bound of

the continuous function $|f'(x)|$ and let the subdivision (a_r) be chosen so that $(a_{r+1}-a_r) < 1/10^n M$. Then

$$f(a_r)-f(x)+(x-a_r)f'(c_r) = 0(n)+0(n),$$

since $\qquad |f'(c_r)| < M \quad \text{and} \quad x-a_r < a_{r+1}-a_r,$

which proves that the polygon approaches the curve $y = f(x)$ as closely as we please. Accordingly we take the length of the arc of $y = f(x)$ to be the limit of the length of the polygon, which is $\int_a^b \sqrt{[1+\{f'(x)\}^2]}\, dx$, as in 10.41.

14.5. Approximations to the value of a definite integral

We have seen that if $a = a_0, a_1,..., a_{p+1} = b$ is an n-chain of the function $f(x)$ then $f(a_0)(a_1-a_0)+f(a_1)(a_2-a_1)+...+f(a_p)(a_{p+1}-a_p)$ differs from $\int_a^b f(x)\, dx$ by $(b-a)0(n-2)$, and so:

14.51. If we divide the interval (a, b) into sufficiently small sub-intervals by the points $a_0, a_1, a_2..., a_{p+1}$, the sum

$$\sum_{r=0}^{p} f(a_r)(a_{r+1}-a_r)$$

is an approximation to the value of $\int_a^b f(x)\, dx$.

14.52. If $f(x)$ and $\phi(x)$ are continuous in (a, b) and

$$f(x)-\phi(x) = 0(n)$$

then $\int_a^b \phi(x)\, dx$ is an approximation to $\int_a^b f(x)\, dx$, with error $(b-a)0(n)$.

For

$$\int_a^b f(x)\, dx - \int_a^b \phi(x)\, dx = \int_a^b \{f(x)-\phi(x)\}\, dx = (b-a)0(n).$$

14.53. If $f(x)$ is differentiable four times then

$$\tfrac{1}{3}h\{f(-h)+4f(0)+f(h)\}$$

is an approximation to the integral $\int_{-h}^{h} f(x)\, dx$ with error $h^5 f^{\text{iv}}(\alpha)/90$, for a certain α in $(-h, h)$.

We observe first that if $\phi(x)$ is a polynomial of the third degree then

$$\int_{-h}^{h} \phi(x)\,dx = \tfrac{1}{3}h\{\phi(-h)+4\phi(0)+\phi(h)\} \quad \textit{exactly}.$$

For if
$$\phi(x) = a+bx+cx^2+dx^3$$
then $\quad \phi(-h)+4\phi(0)+\phi(h) = 4a+2(a+ch^2) = 6a+2ch^2$

and so

$$\int_{-h}^{h} \phi(x)\,dx = 2ah+\tfrac{2}{3}ch^3 = \tfrac{1}{3}h(6a+2ch^2) = \tfrac{1}{3}h\{\phi(-h)+4\phi(0)+\phi(h)\}.$$

Let $R(h)$ denote the difference between

$$\int_{-h}^{h} f(x)\,dx \quad \text{and} \quad \tfrac{1}{3}h\{f(-h)+4f(0)+f(h)\},$$

then

$$R'(h) = f(h)+f(-h)-\tfrac{1}{3}\{f(h)+f(-h)+4f(0)\}-\tfrac{1}{3}h\{f'(h)-f'(-h)\}$$
$$= \tfrac{2}{3}\{f(h)+f(-h)-2f(0)\}-\tfrac{1}{3}h\{f'(h)-f'(-h)\},$$

and $\quad R''(h) = \tfrac{1}{3}\{f'(h)-f'(-h)\}-\tfrac{1}{3}h\{f''(h)+f''(-h)\},$

and $\quad R'''(h) = -\tfrac{1}{3}h\{f'''(h)-f'''(-h)\} = -\tfrac{2}{3}h^2f^{\mathrm{iv}}(\beta_h),$

where $-h<\beta_h<h$, by the mean-value theorem, and β_h is a function of h. Thus $R(0) = R'(0) = R''(0) = R'''(0) = 0$, and so, by 12.52, there is a point c in $[0, h]$ such that

$$\frac{R(h)}{h^5} = \frac{R'''(c)}{5.4.3c^2} = -\frac{1}{90}f^{\mathrm{iv}}(\beta_c).$$

Thus
$$R(h) = -\frac{1}{90}h^5f^{\mathrm{iv}}(\alpha),$$

where $\alpha = \beta_c$, a point in $(-c, c)$, and therefore in $(-h, h)$, which completes the proof.

14.54. Writing $y = x+\dfrac{a+b}{2}$ and $h = \dfrac{b-a}{2}$ so that $y = b$, $\dfrac{a+b}{2}$, and a when $x = h$, 0, and $-h$ respectively, we have

$$\int_a^b f\left(y-\frac{a+b}{2}\right)dy$$
$$= \frac{b-a}{6}\left\{f\left(\frac{b-a}{2}\right)+4f(0)+f\left(\frac{a-b}{2}\right)\right\}-\frac{1}{2880}(b-a)^5f^{\mathrm{iv}}(\alpha),$$

and therefore if $\phi(x) = f\left(x-\frac{a+b}{2}\right)$ so that $f(x) = \phi\left(x+\frac{a+b}{2}\right)$, $f\left(\frac{b-a}{2}\right) = \phi(b)$, $f(0) = \phi\left(\frac{a+b}{2}\right)$, $f\left(\frac{a-b}{2}\right) = \phi(a)$, and

$$f^{\text{iv}}(\alpha) = \phi^{\text{iv}}\left(\alpha+\frac{a+b}{2}\right) = \phi^{\text{iv}}(\rho),$$

say, where $\rho = \alpha+\frac{a+b}{2}$ so that ρ lies between a and b, then

$$\int_a^b \phi(x)\,dx = \frac{b-a}{6}\left\{\phi(a)+4\phi\left(\frac{a+b}{2}\right)+\phi(b)\right\}-\frac{1}{2880}(b-a)^5\phi^{\text{iv}}(\rho).$$

Thus $\frac{b-a}{6}\left\{\phi(a)+4\phi\left(\frac{a+b}{2}\right)+\phi(b)\right\}$ is an approximation to the integral $\int_a^b \phi(x)\,dx$ with error $-\frac{1}{2880}(b-a)^5\phi^{\text{iv}}(\rho)$.

This approximation is known as *Simpson's rule.*

Since the error in the approximation given by Simpson's rule has a bound which is proportional to $(b-a)^5$ (for $\phi^{\text{iv}}(x)$ is bounded) the error of the approximation may be reduced to any assigned amount by subdividing the range of integration and applying Simpson's rule to the integral over each subdivision.

EXAMPLES. (i) Since $\frac{1}{1+x^2}-\frac{1}{1+x^2+x^{10}} = \frac{x^{10}}{(1+x^2)(1+x^2+x^{10})}$, if x lies in $(0, \frac{1}{2})$ then $\frac{1}{1+x^2}-\frac{1}{1+x^2+x^{10}} \leqslant \frac{1}{2^{10}}$, and therefore, by 14.52,

$$\int_0^{\frac{1}{2}} \frac{1}{1+x^2+x^{10}}\,dx = \int_0^{\frac{1}{2}} \frac{1}{1+x^2}\,dx = \tan^{-1}\tfrac{1}{2},$$

with an error less than $\frac{1}{2^{11}}$.

(ii) By Taylor's theorem, with the Cauchy remainder, there is a θ in $[0, 1]$ such that $(1-x)^{-\frac{1}{2}} = 1+\frac{1}{2}x+\frac{3}{4}x^2(1-\theta)/(1-\theta x)^{\frac{5}{2}}$.

Now $\frac{x^2(1-\theta)}{(1-\theta x)^{\frac{5}{2}}} = x^2\frac{1-\theta}{1-\theta x}\left(\frac{1}{1-\theta x}\right)^{\frac{3}{2}}$ and so, if $0 \leqslant x \leqslant \lambda < 1$, we have $\frac{1-\theta}{1-\theta x} < 1$ and $\left(\frac{1}{1-\theta x}\right)^{\frac{3}{2}} < \left(\frac{1}{1-\lambda}\right)^{\frac{3}{2}}$ and so

$$(1-x)^{-\frac{1}{2}} = 1+\tfrac{1}{2}x \quad \text{with error less than} \quad \lambda^2/(1-\lambda)^{\frac{3}{2}}.$$

Hence

$$\int\limits_0^{\frac{1}{2}\pi} \frac{1}{\sqrt{(1-\cdot 19\sin^2\theta)}}\,d\theta = \int\limits_0^{\frac{1}{2}\pi} (1+\cdot 095\sin^2\theta)\,d\theta = 1\cdot 0475\,.\tfrac{1}{2}\pi$$

$$= 1\cdot 6454\ldots,$$

with an error less than $\frac{1}{2}\pi\dfrac{\cdot 19^2}{(1-\cdot 19)^{\frac{3}{2}}} = \frac{1}{2}\pi\dfrac{\cdot 19^2}{\cdot 9^3} < \dfrac{\pi}{40} < \cdot 08$ in defect, so that the value of the integral lies between $1\cdot 645$ and $1\cdot 726$.

(iii) To approximate to $\log 2$.

By Simpson's rule

$$\int\limits_0^1 \frac{1}{1+x}\,dx = \frac{1}{6}\left\{1+\frac{1}{2}+\frac{4}{1+\frac{1}{2}}\right\} = \frac{25}{36} \quad \cdot 6944\ldots,$$

with an error, by excess, of amount

$$\frac{1}{2880}\left\{\frac{d^4}{dx^4}\left(\frac{1}{1+x}\right)\right\}_{x=} \qquad \frac{1}{120}\,\frac{1}{(1+\alpha)^5} < \frac{1}{120} \quad \cdot 008\dot{3},$$

since α lies in $[0, 1]$.

But $\displaystyle\int\limits_0^1 \frac{1}{1+x}\,dx = [\log(1+x)]_0^1 = \log 2$, and so $\log 2$ lies between $\cdot 686$ and $\cdot 69\dot{4}$, so that $\log 2 = \cdot 69$ correct to two decimal places.

Dividing the range $(0, 1)$ into two equal parts, we have

$$\int\limits_0^1 \frac{1}{1+x}\,dx = \int\limits_0^{\frac{1}{2}} \frac{1}{1+x}\,dx + \int\limits_{\frac{1}{2}}^1 \frac{1}{1+x}\,dx$$

$$= \frac{1}{6}\cdot\frac{1}{2}\left(\frac{1}{1}+\frac{1}{1+\frac{1}{2}}+\frac{4}{1+\frac{1}{4}}\right)+\frac{1}{6}\cdot\frac{1}{2}\left(\frac{1}{1+\frac{1}{2}}+\frac{1}{1+1}+\frac{4}{1+\frac{3}{4}}\right)$$

$$= \frac{1747}{2520} = \cdot 69328\ldots,$$

with an error, by excess, less than

$$\frac{1}{120}\cdot\frac{1}{2^5}+\frac{1}{120}\cdot\frac{1}{2^5}\cdot\frac{1}{(1+\frac{1}{2})^5} = \frac{1}{120}\left(\frac{1}{2^5}+\frac{1}{3^5}\right) < \frac{1}{120}\cdot\frac{1}{2^4} = \cdot 00052\ldots,$$

and so $\log 2$ lies between $\cdot 69276$ and $\cdot 69328$, giving $\log 2 = \cdot 693$ correct to three places.

(iv) *Divide the interval $(0, 1)$ into n equal parts by the points r/n, $r = 1, 2,\ldots, n-1$; if n is great enough, these points constitute*

an m-chain of the continuous function x^p, $p > 0$, and therefore, by 14.2, if

$$S_n = \frac{1}{n^p}\left(\frac{2}{n}-\frac{1}{n}\right)+\left(\frac{2}{n}\right)^p\left(\frac{3}{n}-\frac{2}{n}\right)+\ldots+\left(\frac{r}{n}\right)^p\left(\frac{r+1}{n}-\frac{r}{n}\right)+\ldots+$$
$$+\left(\frac{n-1}{n}\right)^p\left(\frac{n}{n}-\frac{n-1}{n}\right),$$

then
$$S_n \to \int\limits^1 x^p\,dx = \frac{}{p+1}.$$

But $S_n = \{1^p+2^p+3^p+\ldots+(n-1)^p\}/n^{p+1}$, and so for large values of n, $1^p+2^p+\ldots+(n-1)^p$ is nearly equal to $\dfrac{n^{p+1}}{p+1}$. Notice, however, that we are here using the expression 'nearly equal' in a different sense from before. In the former sense 'a_n is nearly equal to b_n' means $a_n-b_n \to 0$, but in the present instance 'a_n is nearly equal to b_n' means $a_n/b_n \to 1$. In this latter case it is customary to say that a_n and b_n are *asymptotically* equal. Of course $a_n/b_n \to 1$ does not entail $a_n-b_n \to 0$. In fact it may happen that $a_n/b_n \to 1$ and yet a_n-b_n increases with n. For instance, $\log n$ is arbitrarily great, for great enough n, since $\log n > N$ if $n > e^N$, and $(\log n)/n \to 0$, because $(\log n)/n = M/e^M < 2/M$, where $n = e^M$, and therefore $(n-\log n)/n \to 1$ but $n-(n-\log n) = \log n$.

14.6. If, for any n, $f_n(x)$ is continuous in (a, b) and if
$$\phi(x) = \lim_{(a,b)} f_n(x),$$
then
$$\lim \int\limits^b f_n(x)\,dx = \int\limits^b \phi(x)\,dx.$$

For, by 13.47, $\phi(x)$ is continuous in (a, b) and so, for any n, $f_n(x)-\phi(x)$ is continuous in (a, b) and equals $0(p)$ if n exceeds N_p. Hence $\int\limits_a^b \{f_n(x)-\phi(x)\}\,dx = (b-a).0(p)$ if $n \geqslant N_p$, which proves 14.6.

14.61. If, for each n, $f_n(x)$ is differentiable and $f_n(x)$ tends to a limit $f(x)$ in (a, b), and if $f_n'(x)$ is interval-convergent in (a, b), then $f(x)$ is differentiable and $f_n'(x) \to f'(x)$.

Let $\phi(x)$ be the limit in (a, b) of the convergent sequence $f'_n(x)$; then, by 14.6,

$$\int_a^t f'_n(x)\,dx \to \int_a^t \phi(x)\,dx.$$

But $\int_a^t f'_n(x)\,dx = f_n(t) - f_n(a)$, and $f_n(x) \to f(x)$, and therefore

$$f(t) - f(a) = \int_a^t \phi(x)\,dx,$$

which proves that $f(t)$ is differentiable and $f'(t) = \phi(t)$.

XV

PARTIAL DIFFERENTIATION

FUNCTIONS OF SEVERAL VARIABLES. DOUBLE AND REPEATED LIMITS. CONTINUITY AND DERIVABILITY. THE TAYLOR SERIES. EULER'S THEOREM. INVERSION OF A FUNCTIONAL RELATION. PARTIAL DERIVATIVE OF AN INTEGRAL

15. In considering the properties of functions of two variables we shall have frequent occasion to specify the range in which the variables lie; instead of saying of a pair of variables x, y that x lies in a certain interval (a, b) and y lies in (c, d) we shall say, briefly, that the *point* (x, y) lies in the *rectangle* $(a, b)(c, d)$, or when it is unnecessary to be more explicit, just that (x, y) lies in a rectangle R. The points (a, c), (a, d), (b, c), (b, d) are called the vertices of the rectangle $(a, b)(c, d)$ and $\left(\frac{a+b}{2}, \frac{c+d}{2}\right)$ its centre. Observe, however, that nothing more is implied by the terminology than we have stated; to say that the point (x, y) lies in the rectangle $(a, b)(c, d)$ means only that $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$, and we shall also extend the use of the term *interval* to cover the notion of a pair of numbers lying in a certain pair of intervals. We have introduced the new term 'rectangle' here, however, to accentuate certain important differences between the theory of functions of one variable and functions of two variables.

15.1. *If* $f(X, y)-\phi(x, y) = O(p)$ for any points (x, y), (X, y) in a rectangle R, such that $X-x = O(q)$, where q depends *only* upon p (and not upon x, X, or y), and $X \neq x$, then we say that $\phi(x, y)$ is the *interval-limit* of $f(X, y)$ as X *tends to* x, and we write

$$\lim_{X \to x(R)} f(X, y) = \phi(x, y).$$

15.11. Similarly, *if* $f(x, Y)-\psi(x, y) = O(p)$ for any points (x, y), (x, Y) in a rectangle R, such that $Y-y = O(q)$, q depending on p alone, and $Y \neq y$, then we say that $\psi(x, y)$ is the interval-limit of $f(x, Y)$ as Y tends to y, and we write

$$\lim_{Y \to y(R)} f(x, Y) = \psi(x, y).$$

15.12. *If* $f(X, Y)-\theta(x, y) = 0(p)$ for any (x, y), (X, Y) in R such that $X-x$ and $Y-y$ are $0(q)$ (not both zero), q depending *only* on p, then $\theta(x, y)$ is called the (interval) *double limit* of $f(X, Y)$ as X tends to x and Y tends to y, and we write

$$\lim_{X, Y\to x, y(R)} f(X, Y) = \theta(x, y).$$

15.13. *If* $\lim\limits_{X\to x(R)} f(X, Y) = \phi(x, Y)$, in R, and *if*

$$\lim_{Y\to y(R)} \phi(x, Y) = \lambda(x, y),$$

in R, then $\lambda(x, y)$ is called the (interval) *repeated limit* of $f(X, Y)$, as X tends to x and Y tends to y, and we write

$$\lim_{Y\to y}\lim_{X\to x(R)} f(X, Y) = \lambda(x, y).$$

If $\lim\limits_{Y\to y(R)} f(X, Y) = \psi(X, y)$, in R, and *if*

$$\lim_{X\to x(R)} \psi(X, y) = \mu(x, y),$$

in R, then $\mu(x, y)$ is called the (interval) *repeated limit* of $f(X, Y)$, as Y tends to y and X tends to x, and we write

$$\lim_{X\to x}\lim_{Y\to y(R)} f(X, Y) = \mu(x, y).$$

15.14. The interval double limit, and the two interval repeated limits of f(X, Y), all exist and are equal provided only the two single interval-limits $\lim\limits_{X\to x(R)} f(X,y)$ and $\lim\limits_{Y\to y(R)} f(x,Y)$ exist.

Proof. Let $\phi(x, y)$, $\psi(x, y)$ be the values of the limits $\lim\limits_{X\to x(R)} f(X, y)$, $\lim\limits_{Y\to y(R)} f(x, Y)$ in a rectangle R, so that

$$f(X, y)-\phi(x, y) = 0(p), \qquad f(x, Y)-\psi(x, y) = 0(p)$$

for any (x, y), (X, y), and (x, Y), in R, such that $X-x, Y-y = 0(q)$. Hence

$$f(X, Y)-\psi(X, y) = 0(p), \qquad f(X, Y)-\phi(x, Y) = 0(p)$$

and so $$\psi(X, y)-\phi(x, Y) = 0(p-1),$$

whence $$\psi(X, y)-\psi(X^*, y) = 0(p-2)$$

for any X, X^* such that

$$X-x,\ X^*-x \quad 0(q).$$

Hence if x_n is a sequence which converges to x, then for given (x_n) and y, the sequence $\psi(x_n, y)$ is convergent, for

$$\psi(x_m, y)-\psi(x_n, y) = 0(p-2) \quad \text{provided } x-x_n,\ x-x_m = 0(q),$$

i.e. provided m and n are greater than some N, which depends upon p, y and the sequence (x_n). Let λ be the limit of $\psi(x_n, y)$, and let X_n be another sequence which converges to x; then for n greater than or equal to a certain K,

$$\lambda-\psi(x_n, y) = 0(p) \quad \text{and} \quad \psi(x_n, y)-\psi(X_n, y) = 0(p),$$

and therefore $\qquad \lambda-\psi(X_n, y) = 0(p-1),$

which proves that λ is the limit of the sequence $\psi(X_n, y)$. Thus λ does not depend upon the particular sequence (x_n) but only upon x and y, and so $\lambda = \lambda(x, y)$, a function of x and y alone.

Since for each value of x and y, $\lambda(x, y)$ is the limit of $\psi(X_n, y)$ as X_n tends to x, it follows that

$$\psi(X_n, y)-\lambda(x, y) = 0(p) \quad \text{provided } n \geqslant K(p, x, y).$$

Choose an n so great that $x-X_n = 0(q)$ and $n \geqslant K(p, x, y)$, then for *any* X, such that $X-x = 0(q)$, we have

$$\begin{aligned} \psi(X, y)-\lambda(x, y) &= \psi(X, y)-\psi(X_n, y)+\psi(X_n, y)-\lambda(x, y) \\ &= 0(p-2)+0(p) = 0(p-3), \end{aligned}$$

which proves that

$$\lim_{X\to x(R)} \psi(X, y) = \lambda(x, y),$$

i.e. $\qquad \lim\limits_{X\to x}\ \lim\limits_{Y\to y(R)} f(X, Y) = \lambda(x, y).$

In a similar way we can prove that $\lim\limits_{Y\to y(R)} \phi(x, Y)$ exists and has the value $\mu(x, y)$ say; but $\psi(X, y)-\phi(x, Y) = 0(p-1)$ and therefore $\lambda(x, y)-\mu(x, y) = 0(p-4)$, for any p, which proves that

$$\lambda(x, y) = \mu(x, y) \quad \text{in } R.$$

Moreover, if $X-x,\ Y-y = 0(q)$,

$$f(X, Y)-\phi(x, Y) = 0(p), \qquad \phi(x, Y)-\mu(x, y) = 0(p-3)$$

and therefore $\qquad f(X, Y)-\mu(x, y) = 0(p-4),$

which proves that

$$\lim_{X,Y\to x,y(R)} f(X, Y) = \mu(x, y).$$

Thus the two interval repeated limits and the interval double limit all exist and are equal.

15.2. The function $f(x, y)$ is said to be *uniformly* (or *interval*) *x-continuous in R* if $\lim_{X \to x(R)} f(X, y) = f(x, y)$ in R.

$f(x, y)$ is uniformly (or interval) *y-continuous* in R if

$$\lim_{Y \to y(R)} f(x, Y) = f(x, y) \quad \text{in } R.$$

If $f(x, y)$ is both uniformly x-continuous and y-continuous in R then $f(x, y)$ is said to be uniformly *continuous* in R.

For brevity we shall contract the expression 'uniformly continuous in R' to 'continuous in R'.

A necessary and sufficient condition for continuity is

$$\lim_{X,Y \to x,y(R)} f(X, Y) = f(x, y),$$

for if $\lim_{X \to x(R)} f(X, Y) = f(x, Y)$ and $\lim_{Y \to y(R)} f(x, Y) = f(x, y)$ then $\lim_{Y \to y} \lim_{X \to x(R)} f(X, Y) = f(x, y)$, and so, by 15.14,

$$\lim_{X,Y \to x,y(R)} f(X, Y) = f(x, y);$$

conversely, if $\lim_{X,Y \to x,y(R)} f(X, Y) = f(x, y)$ then, taking $Y = y$, we have $\lim_{X \to x(R)} f(X, y) = f(x, y)$, and taking $X = x$,

$$\lim_{Y \to y(R)} f(x, Y) = f(x, y),$$

so that $f(x, y)$ is continuous.

15.201. If $x_r = a + r(b-a)/n$, $y_s = c + s(d-c)/m$, then the set of intervals (rectangles)

$$(x_r, x_{r+1})(y_s, y_{s+1}), \quad \begin{array}{l} r = 0, 1, 2, \ldots, n-1, \\ s = 0, 1, 2, \ldots, m-1, \end{array}$$

is called a subdivision of the interval $(a, b)(c, d)$ into mn equal sub-intervals.

15.202. If $f(x, y)$ is continuous in the interval $(a, b)(c, d)$ then we can divide $(a, b)(c, d)$ into a finite number of sub-intervals such that

$$f(x, y) - f(X, Y) = 0(p)$$

for any two points (x, y), (X, Y) in the same sub-interval.

For $f(x, y) - f(X, Y) = 0(p)$ if $X - x$, $Y - y = 0(q)$. Choose n so that $(b-a)/n$, $(d-c)/n = 0(q)$ and divide $(a, b)(c, d)$ into n^2 equal sub-intervals. If $(x_r, x_{r+1})(y_s, y_{s+1})$ is any one of the sub-intervals,

then $x_{r+1}-x_r$, $y_{s+1}-y_s = 0(q)$, and so if (x, y), (X, Y) are any two points in $(x_r, x_{r+1})(y_s, y_{s+1})$, then $X-x$, $Y-y = 0(q)$ and therefore $f(x, y)-f(X, Y) = 0(p)$.

15.21. If $f(x, y)$ is continuous in R then it is bounded in R.

Divide R into n^2 equal intervals such that $f(x, y)-f(X, Y) = 0(1)$ for any two points (x, y), (X, Y) of the same sub-interval.

Let (x, y) be any point of R; then (x, y) falls into one of the sub-intervals $(x_r, x_{r+1})(y_s, y_{s+1})$, say ($(x, y)$ may in fact be a vertex of four of the sub-intervals, in which case we may choose any one of the four as the sub-interval in which (x, y) lies). Then

$$\begin{aligned} f(x, y) = \{f(x, y)-f(x_r, y_s)\}+\{f(x_r, y_s)-f(x_{r-1}, y_s)\}+\ldots+ \\ +\{f(x_1, y_s)-f(x_0, y_s)\}+\{f(x_0, y_s)-f(x_0, y_{s-1})\}+ \\ +\{f(x_0, y_{s-1})-f(x_0, y_{s-2})\}+\ldots+\{f(x_0, y_1)-f(x_0, y_0)\}+f(x_0, y_0) \end{aligned}$$

and so

$$|f(x, y)-f(x_0, y_0)| = (r+s+1)0(1) < (2n+1)/10,$$

which proves that $f(x, y)$ is bounded in R.

15.22. If $f(x, y)$ is continuous, and $\phi(t)$, $\psi(t)$ are continuous, then $f(\phi(t), \psi(t))$ is continuous.

For $\phi(T)-\phi(t) = 0(p)$, $\psi(T)-\psi(t) = 0(p)$ when $T-t = 0(q)$ and $f(X, Y)-f(x, y) = 0(r)$ when $X-x$, $Y-y = 0(p)$, so that

$$f(\phi(T), \psi(T))-f(\phi(t), \psi(t)) = 0(r),$$

which proves that $f(\phi(t), \psi(t))$ is continuous.

15.23. If $f(x, y)$ is continuous then $f(x, y)$ takes any value between any two of its values.

Let V_1 and V_2 be the values of $f(x, y)$ at the points (x_1, y_1) and (x_2, y_2), and let $\phi(t)$, $\psi(t)$ be two linear functions which take the values x_1, y_1 and x_2, y_2 for $t = t_1$ and $t = t_2$ respectively. Then $f(\phi(t), \psi(t))$ is a continuous function which takes the values V_1, V_2 at the points t_1, t_2 and so takes any value between V_1 and V_2 at some point $\phi(t)$, $\psi(t)$ for a t between t_1 and t_2.

15.24. If $f(x, y)$ is continuous in R and $|f(x, y)| \geqslant \alpha > 0$ in R, then $f(x, y)$ is of constant sign in R.

For if (x_1, y_1), (x_2, y_2) are any two points in R, then $f(x, y)$ takes any value between $f(x_1, y_1)$ and $f(x_2, y_2)$. But since $|f(x, y)| \geqslant \alpha$, zero is not a value of $f(x, y)$ and so zero does not lie between

$f(x_1, y_1)$ and $f(x_2, y_2)$, which proves that $f(x_1, y_1)$ and $f(x_2, y_2)$ are of the same sign.

15.25. The functions $f(x, y)$, $g(x, y)$ are continuous in R and at each point of R at least one of the numbers $|f(x, y)|$, $|g(x, y)|$ exceeds δ. Then we can divide R into a finite number of sub-intervals such that in each sub-interval, ρ, *either* $|f(x, y)| > \frac{1}{2}\delta$ at all points of ρ, *or* $|g(x, y)| > \frac{1}{2}\delta$ at all points of ρ.

Proof. By 15.202, we can divide R into a finite number of sub-intervals such that in each sub-interval ρ,

$$|f(x, y)-f(X, Y)| < \tfrac{1}{2}\delta, \qquad |g(x, y)-g(X, Y)| < \tfrac{1}{2}\delta,$$

for any (x, y), (X, Y) in ρ. Let X, Y be the centre of ρ. Either $|f(X, Y)|$ or $|g(X, Y)|$ exceeds δ; suppose the former, then

$$|f(x, y)| = |\{f(x, y)-f(X, Y)\}+f(X, Y)| > \delta-\tfrac{1}{2}\delta = \tfrac{1}{2}\delta,$$

for all points (x, y) in ρ. Similarly, if $|g(X, Y)| > \delta$, then

$$|g(x, y)| > \tfrac{1}{2}\delta$$

at all points of ρ.

15.3. A function $f(x, y)$ is said to be *uniformly x-differentiable* in R if $f(x, y)$ is x-continuous in R and if we can determine a function $\phi(x, y)$ and a function $q(p)$ such that (for different x, X)

$$\{f(X, y)-f(x, y)\}/(X-x) = \phi(x, y)+0(p)$$

for any (x, y), (X, y) satisfying $X-x = 0(q)$, q depending only upon p. The function $\phi(x, y)$ is called the *uniform x-derivative* of $f(x, y)$. $f(x, y)$ is said to be *uniformly y-differentiable* in R if $f(x, y)$ is y-continuous in R, and if we can determine a function $\psi(x, y)$ and a function $q(p)$ such that $\{f(x, Y)-f(x, y)\}/(Y-y) = \psi(x, y)+0(p)$ for any different (x, y), (x, Y) in R satisfying $Y-y = 0(q)$. $\psi(x, y)$ is called the *uniform y-derivative* of $f(x, y)$.

Observe that for any constant value of y, the x-derivative of $f(x, y)$ is just the derivative of $f(x, y)$ as a function of the single variable x, and for a constant value of x, the y-derivative is just the derivative of $f(x, y)$ as a function in the single variable y. Thus in forming the x- or y-derivative of a function $f(x, y)$ we may employ the technique already established for differentiating $f(x, y)$, treating x as a mere constant or y as a mere constant as the case may be.

We shall generally denote the x- and y-derivatives of $f(x, y)$

by $f_x(x, y)$ and $f_y(x, y)$ respectively; another common notation is $\frac{\partial}{\partial x}f(x, y)$ and $\frac{\partial}{\partial y}f(x, y)$. As we have already observed, the x- and y-derivatives are ordinary derivatives formed by regarding y as a constant and x as a constant respectively, and could therefore be denoted by $\frac{d}{dx}f(x, y)$ and $\frac{d}{dy}f(x, y)$; it is customary, however, to reserve the notation $\frac{d}{dx}f(x, y)$ for the derivative of $f(x, y)$ when y is regarded, not as a constant, but as a function of x, and the [illegible] $\frac{d}{dy}f(x, y)$ for the derivative of $f(x, y)$ when x is regarded as a function of y. It is for this reason that the curly '∂' is used, instead of the straight 'd' to denote the x- and y-derivatives of a function of two independent variables, and not because there is something new in the kind of differentiation introduced in this section.

15.301. It is useful to maintain a distinction between the signs $\frac{\partial f}{\partial x}$ and f_x and between the signs $\frac{\partial f}{\partial y}$ and f_y in the case when the argument places of f are filled by functions. Thus $\frac{\partial}{\partial x}f\{g(x, y), h(x, y)\}$ is the x-derivative of $\phi(x, y)$, where $\phi(x, y) = f\{g(x, y), h(x, y)\}$, but $f_x\{g(x, y), h(x, y)\}$ is the *result* of substituting $g(x, y)$ for x and $h(x, y)$ for y in the function $f_x(x, y)$. Thus for instance

$$f_x(y, x) = \frac{\partial}{\partial y}f(y, x)$$

for $f_x(y, x)$ is the result of differentiating $f(x, y)$ with respect to x and *then* interchanging x and y, which amounts to differentiating $f(y, x)$ with respect to y.

In forming $\frac{\partial}{\partial x}f\{g(x, y), h(x, y)\}$ the variables x and y in $f(x, y)$ are replaced by $g(x, y)$ and $h(x, y)$ *before* differentiation, and *after* differentiation in forming $f_x\{g(x, y), h(x, y)\}$.

The difference between $\frac{\partial f}{\partial x}$ and f_x is analogous to that between $\frac{df}{dx}$ and $f'(x)$ for functions of a single variable.

15.302. If $f(x, y)$ is both uniformly x-differentiable and uniformly y-differentiable in R, then $f(x, y)$ is said to be uniformly differentiable in R.

For brevity we shall generally write 'differentiable in R' for 'uniformly differentiable in R'.

15.31. If $f(x, y)$ is uniformly differentiable in R then both the x-derivative and the y-derivative are continuous in R.

Proof. Since $f(x, y)$ is both x- and y-differentiable, therefore $f(x, y)$ is both x- and y-continuous, i.e. $f(x, y)$ is continuous.

Now $$\frac{f(X, y)-f(x, y)}{X-x} = f_x(x, y)+0(p)$$

for any (x, y), (X, y) in R such that $X-x = 0(q)$; interchanging x, X we find

$$\frac{f(x, y)-f(X, y)}{x-X} = f_x(X, y)+0(p)$$

and so $$f_x(X, y)-f_x(x, y) = 0(p-1),$$

which proves that $f_x(x, y)$ is x-continuous in R.

Let $X-x = 1/10^{q+1}$; since $f(x, y)$ is continuous, in R, we can choose r so that $f(x, Y)-f(x, y) = 0(p+q+1)$, for any (x, y), (x, Y) such that $Y-y = 0(r)$. Hence

$$\begin{aligned} f_x(x, Y)-f_x(x, y) & \\ &= \frac{f(X, Y)-f(X, y)-\{f(x, Y)-f(x, y)\}}{X-x}+0(p)-0(p), \\ &= 10^{q+1}\{0(p+q+1)-0(p+q+1)\}+0(p)+0(p) \\ &= 0(p)+0(p)+0(p)+0(p) = 0(p-1), \end{aligned}$$

provided only $Y-y = 0(r)$, r depending on p alone, which proves that $f_x(x, y)$ is y-continuous in R. Thus $f_x(x, y)$ is continuous in R. Similarly, $f_y(x, y)$ is continuous in R.

15.32. If $f(x, y)$ is differentiable in R then

$$\mathbf{f}(\mathbf{X}, \mathbf{Y})-\mathbf{f}(\mathbf{x}, \mathbf{y}) = (\mathbf{X}-\mathbf{x})\mathbf{f}_\mathbf{x}(\mathbf{x}, \mathbf{y})+(\mathbf{Y}-\mathbf{y})\mathbf{f}_\mathbf{y}(\mathbf{x}, \mathbf{y})+ \\ +(\mathbf{X}-\mathbf{x})0(\mathbf{p})+(\mathbf{Y}-\mathbf{y})0(\mathbf{p})$$

for any (x, y), (X, Y) such that $X-x$, $Y-y = 0(q)$.

For

$$f(X,Y)-f(x,y)-(X-x)f_x(x,y)-(Y-y)f_y(x,y)$$
$$= f(X,Y)-f(x,Y)+$$
$$+f(x,Y)-f(x,y)-(X-x)f_x(x,y)-(Y-y)f_y(x,y)$$
$$= (X-x)\{f_x(x,Y)-f_x(x,y)+0(p)\}+$$
$$+(Y-y)\{f_y(x,y)-f_y(x,y)+0(p)\}$$
$$= (X-x)\{0(p)+0(p)\}+(Y-y)0(p),$$

since $f_x(x,y)$ is y-continuous,

$$= (X-x)0(p-1)+(Y-y)0(p-1).$$

15.33. If $f(x,y)$ is continuous in R and if

$$f(X,Y)-f(x,y)$$
$$= (X-x)\phi(x,y)+(Y-y)\psi(x,y)+(X-x)0(p)+(Y-y)0(p)$$

for any (x,y), (X,Y) such that $X-x$, $Y-y = 0(q)$, then $f(x,y)$ is differentiable in R and $\phi(x,y)$, $\psi(x,y)$ are the x- and y-derivatives of $f(x,y)$.

For taking $Y = y$ we have

$$\frac{f(X,y)-f(x,y)}{X-x} = \phi(x,y)+0(p)$$

and taking $X = x$,

$$\frac{f(x,Y)-f(x,y)}{Y-y} = \psi(x,y)+0(p),$$

proving that $f(x,y)$ is both x- and y-derivable with derivatives $\phi(x,y)$ and $\psi(x,y)$.

15.4. Observe that, unlike the case for a single variable, a differentiable function $f(x,y)$ is continuous by *definition*; in the case of a function of a single variable the continuity of a differentiable function is a consequence of the provable continuity of the uniform derivative.

The two-variable function can be put on the same level as the one-variable function, if we formulate the definition of differentiability in the following way:

$f(x,y)$ is uniformly x-differentiable in R if (for different x, X)

$$\frac{f(X,Y)-f(x,Y)}{X-x} = \phi(x,y)+0(p)$$

for any (x, Y), (X, Y) in R such that $X-x$, $Y-y = 0(q)$, where q depends on p *alone*.

$f(x, y)$ is uniformly y-differentiable in R if (for different y, Y)

$$\frac{f(X, Y)-f(X, y)}{Y-y} = \psi(x, y)+0(p)$$

for any (X, y), (X, Y) in R such that $X-x$, $Y-y = 0(q)$, q depending on p *alone*.

Denoting the derivatives by $f_x(x, y)$, $f_y(x, y)$ as before, it follows that

$$\begin{aligned} f(X, Y)&-f(x, y)-(X-x)f_x(x, y)-(Y-y)f_y(x, y) \\ &= f(X, Y)-f(x, Y)+ \\ &\qquad +f(x, Y)-f(x, y)-(X-x)f_x(x, y)-(Y-y)f_y(x, y) \\ &= (X-x)\{f_x(x, y)-f_x(x, y)+0(p)\}+ \\ &\qquad +(Y-y)\{f_y(x, y)-f_y(x, y)+0(p)\} \\ &= (X-x)0(p)+(Y-y)0(p). \end{aligned}$$

15.401. Note that, in § 15 3, the condition which defines the x-derivative contains three variables, x, y, and Y, whereas the condition in § 15.4 contains four variables x, y, X, and Y. Thus the second condition on $f(x, y)$ is far more stringent than the first, and it is to be expected that we can learn more from it about the behaviour of $f(x, y)$.

15.41. The derivatives $f_x(x, y)$ and $f_y(x, y)$, as defined in 16.4, are continuous functions.

The proof that $f_x(x, y)$ is x-continuous proceeds exactly as in 15.31. Taking $y = Y$ in the defining equation 15.4 and subtracting from the resulting equation, 15.4 itself, we have

$$f_x(x, Y)-f_x(x, y) = 0(p-1),$$

since the left-hand side of 15.4 is unchanged by replacing y by Y, and so $f_x(x, y)$ is y-continuous. Thus $f_x(x, y)$ is continuous, and similarly $f_y(x, y)$ is continuous.

15.42. If $f(x, y)$ is differentiable in R, then $f(x, y)$ is continuous in R.

Replace Y by y in 15.4 and we have

$$\begin{aligned} f(X, y)-f(x, y) &= (X-x)f_x(x, y)+(X-x)0(p) \\ &= 0(r) \end{aligned}$$

if $X-x$ is small enough, since $f_x(x, y)$ is continuous and therefore bounded.

Thus $f(x, y)$ is x-continuous.

Similarly, since $f(x, Y)-f(x, y) = (Y-y)f_y(x, y)+(Y-y)\,0(p)$ and $f_y(x, y)$ is bounded, therefore $f(x, y)$ is y-continuous. Hence $f(x, y)$ is continuous.

15.43. The definitions 15.3 and 15.4 are equivalent

15.4 implies 15.3 because 15.4 implies $f(x, y)$ is continuous (by 15.42) and we may take $Y = y$ in 15.4. Furthermore 15.3 implies 15.4; for if $f(x, y)$ is continuous and

$$\frac{f(X, y)-f(x, y)}{X-x} = f_x(x, y)+0(p)$$

then by 15.31 $f_x(x, y)$ is continuous, and therefore

$$\begin{aligned}\frac{f(X, Y)-f(x, Y)}{X-x} &= f_x(x, Y)+0(p)\\ &= f_x(x, y)+0(p)+0(p)\\ &= f_x(x, y)+0(p-1)\end{aligned}$$

provided only $Y-y = 0(r)$ for an r depending on p alone.

Similarly, $\dfrac{f(X, Y)-f(X, y)}{Y-y} = f_y(x, y)+0(p-1)$,

which proves that 15.4 follows from 15.3.

15.44. Repeated differentiation

If the x- and y-derivatives of $f(x, y)$ are themselves differentiable functions, $f(x, y)$ is said to be differentiable twice. The x-derivatives of $f_x(x, y)$ and $f_y(x, y)$ are denoted by $f_{xx}(x, y)$ and $f_{yx}(x, y)$ respectively, and the y-derivatives by $f_{xy}(x, y)$ and $f_{yy}(x, y)$ respectively. In the '∂' notation the x-derivatives of $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are written $\dfrac{\partial^2 f}{\partial x^2}$ and $\dfrac{\partial^2 f}{\partial x \partial y}$ respectively, and the y-derivatives are written $\dfrac{\partial^2 f}{\partial y \partial x}$ and $\dfrac{\partial^2 f}{\partial y^2}$ respectively.

If all the second derivatives are differentiable we obtain the third derivatives

$$\frac{\partial^{m+n} f}{\partial x^m \partial y^n} \quad \text{and} \quad \frac{\partial^{m+n} f}{\partial y^n \partial x^m}, \quad m+n = 3, \quad m \geqslant 0, \quad n \geqslant 0,$$

and from these the fourth derivatives are formed, and so on.

The fundamental theorem of this section is that the *order* of differentiation is immaterial, i.e. $\dfrac{\partial^{m+n}f}{\partial x^m \partial y^n} = \dfrac{\partial^{m+n}f}{\partial y^n \partial x^m}$ for any m, n, and this is an immediate consequence of

15.45. If $f_x(x, y)$ and $f_y(x, y)$ are differentiable, then $f_{xy} = f_{yx}$.

Let $\phi(x)$ denote $f(x, y+h)-f(x, y)$ and $\psi(y)$ denote

$$f(x+h, y)-f(x, y).$$

Then since $f(x, y)$ is x-differentiable twice, $\phi'(x)$ and $\phi''(x)$ exist and since $f(x, y)$ is y-differentiable twice, $\psi'(y)$ and $\psi''(y)$ exist. Hence if $h = 0(q)$

$$\phi(x+h)-\phi(x) = h\phi'(x)+\frac{h^2}{2!}\{\phi''(x)+0(p)\}$$

and
$$\psi(y+h)-\psi(y) = h\psi'(y)+\frac{h^2}{2!}\{\psi''(y)+0(p)\}.$$

But $\phi'(x) = f_x(x, y+h)-f_x(x, y) = h\{f_{xy}(x, y)+0(p)\}$, since f_x is y-differentiable, and $\phi''(x) = f_{xx}(x, y+h)-f_{xx}(x, y) = 0(p)$, since f_{xx} is continuous, provided $|h|$ is small enough.

Thus
$$\phi(x+h)-\phi(x) = h^2\{f_{xy}+0(p-1)\}.$$
Similarly
$$\psi(y+h)-\psi(y) = h^2\{f_{yx}+0(p-1)\}.$$
But

$$\begin{aligned}\phi(x+h)-\phi(x) &\\ &= \{f(x+h, y+h)-f(x+h, y)\}-\{f(x, y+h)-f(x, y)\}\\ &= \{f(x+h, y+h)-f(x, y+h)\}-\{f(x+h, y)-f(x, y)\}\\ &= \psi(y+h)-\psi(y),\end{aligned}$$

and therefore $f_{xy} = f_{yx}+0(p-2)$; since this is true for any p it follows that $f_{xy} = f_{yx}$.

The equality of f_{xy} and f_{yx} may also be established by the *mean-value* theorem.

Denoting $f(x, y+k)-f(x, y)$ by $\phi(x)$ and $f(x+h, y)-f(x, y)$ by $\psi(y)$ we have

$$\begin{aligned}\phi(x+h)-\phi(x) &\\ &= h\phi'(x+\theta h)\\ &= h\{f_x(x+\theta h, y+k)-f_x(x+\theta h, y)\}\\ &= hk\{f_{xy}(x+\theta h, y)+0(p)\}, \quad \text{since } f_x \text{ is } y\text{-differentiable},\\ &= hk\{f_{xy}(x, y)+0(p)\}, \quad \text{since } f_{xy} \text{ is continuous.}\end{aligned}$$

Similarly $\quad \psi(y+k)-\psi(y) = hk\{f_{yx}(x, y)+0(p)\};$

but

$$\phi(x+h)-\phi(x) = f(x+h, y+k)-f(x+h, y)-f(x, y+k)+f(x, y)$$
$$= \psi(y+k)-\psi(y)$$

and therefore $f_{xy}(x, y) = f_{yx}(x, y)$.

15.46. The definitions and theorems in §§ 15–15.45 readily extend to functions of three or more variables. For instance a function $f(x, y, z)$ is x-continuous in an interval I if

$$f(X, y, z)-f(x, y, z) = 0(p)$$

for any (x, y, z) and (X, y, z) in I such that $X-x = 0(q)$, and $f(x, y, z)$ is x-differentiable in I if it is x-continuous in I and if there is a function $\phi(x, y, z)$ such that (for different X, x)

$$\frac{f(X, y, z)-f(x, y, z)}{X-x}-\phi(x, y, z) = 0(p)$$

for any (x, y, z), (X, y, z) in I such that $X-x = 0(q)$. Of course $f(x, y, z)$ has *three* derivatives f_x, f_y, f_z and exactly as in 15.45 we can show that $f_{xy} = f_{yx}$ and $f_{xz} = f_{zx}$, etc.

15.5. If $f(x, y)$ is differentiable and if $x(t)$ and $y(t)$ are differentiable functions of t then

$$\frac{df(x, y)}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt},$$

where the variables in $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are supposed replaced by the functions $x(t)$ and $y(t)$ *after* differentiation, and in $df(x, y)/dt$ *before* differentiation.

Proof. Denote $x(t)$, $y(t)$ by x, y and $x(T)$, $y(T)$ by X, Y respectively. Then

$$f(x(T), y(T))-f(x(t), y(t)) = f(X, Y)-f(x, y)$$
$$= (X-x)\{f_x(x, y)+0(p)\}+(Y-y)\{f_y(x, y)+0(p)\}, \quad \text{by 15.32,}$$

provided $X-x$, $Y-y = 0(q)$.

But $x(t)$ and $y(t)$ are differentiable and so

$$\frac{X-x}{T-t} = \frac{x(T)-x(t)}{T-t} = \frac{dx}{dt}+0(p)$$

and
$$\frac{Y-y}{T-t} = \frac{y(T)-y(t)}{T-t} = \frac{dy}{dt}+0(p)$$
provided $T-t = 0(r)$. Since dx/dt and dy/dt are continuous, they are bounded, and so we may choose r so that when $T-t = 0(r)$, then $X-x$, $Y-y = 0(q)$. Hence when $T-t = 0(r)$, we have
$$\frac{f(x(T), y(T))-f(x(t), y(t))}{T-t} - \left(\frac{dx}{dt}f_x+\frac{dy}{dt}f_y\right)$$
$$= \left(f_x+f_y+\frac{dx}{dt}+\frac{dy}{dt}\right)0(p)+0(2p-1).$$
But f_x, f_y, dx/dt, dy/dt being continuous, are bounded, and so
$$\frac{d}{dt}f(x(t), y(t)) = \lim_{T\to t}\frac{f(x(T), y(T))-f(x(t), y(t))}{T-t} = \frac{dx}{dt}f_x+\frac{dy}{dt}f_y,$$
which proves 15.5.

15.51. By means of 15.5 we can readily form the second, third, and so on, derivatives of $f(x, y)$ with respect to t, provided that $f(x, y)$, $x(t)$, and $y(t)$ are differentiable sufficiently often. For instance
$$\frac{d^2f}{dt^2} = \frac{d}{dt}\left(\frac{df}{dt}\right) = \frac{d}{dt}\left(\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}\right)$$
$$= \frac{\partial f}{\partial x}\frac{d^2x}{dt^2}+\frac{\partial f}{\partial y}\frac{d^2y}{dt^2}+\frac{dx}{dt}\frac{d}{dt}\left(\frac{\partial f}{\partial x}\right)+\frac{dy}{dt}\frac{d}{dt}\left(\frac{\partial f}{\partial y}\right).$$
Since formula 15.5 is true for any differentiable function $f(x, y)$ we may take for $f(x, y)$ the differentiable function $\partial f/\partial x$, whence
$$\frac{d}{dt}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}\frac{dx}{dt}+\frac{\partial^2 f}{\partial x\partial y}\frac{dy}{dt},$$
and similarly
$$\frac{d}{dt}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\partial y}\frac{dx}{dt}+\frac{\partial^2 f}{\partial y^2}\frac{dy}{dt}.$$
Thus
$$\mathbf{\frac{d^2f}{dt^2} = \frac{\partial^2 f}{\partial x^2}\left(\frac{dx}{dt}\right)^2+2\frac{\partial^2 f}{\partial x\partial y}\frac{dx}{dt}\frac{dy}{dt}+\frac{\partial^2 f}{\partial y^2}\left(\frac{dy}{dt}\right)^2+\frac{\partial f}{\partial x}\frac{d^2x}{dt^2}+\frac{\partial f}{\partial y}\frac{d^2y}{dt^2}}.$$
This formula may be abbreviated by writing $\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)^2 f$ for the set of terms $\left(\frac{dx}{dt}\right)^2\frac{\partial^2 f}{\partial x^2}+2\frac{dx}{dt}\frac{dy}{dt}\frac{\partial^2 f}{\partial x\partial y}+\left(\frac{dy}{dt}\right)^2\frac{\partial^2 f}{\partial y^2}$, so that in expanding the binomial $\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)^2$ we treat the operators

$\partial/\partial x$, $\partial/\partial y$ like variables. The formula becomes

$$\frac{d^2f}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)^2 f+\frac{\partial f}{\partial x}\frac{d^2x}{dt^2}+\frac{\partial f}{\partial y}\frac{d^2y}{dt^2}.$$

Differentiating again we find

$$\frac{d^3f}{dt^3} = \frac{d}{dt}\left(\frac{d^2f}{dt^2}\right) = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)^2 f+$$
$$+\frac{d^2x}{dt^2}\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)\frac{\partial f}{\partial x}+\frac{d^2y}{dt^2}\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)\frac{\partial f}{\partial y}+$$
$$+\frac{\partial f}{\partial x}\frac{d^3x}{dt^3}+\frac{\partial f}{\partial y}\frac{d^3y}{dt^3},$$

whence

$$\frac{d^3f}{dt^3} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)^3 f+\frac{\partial^2 f}{\partial x^2}\frac{d^2x}{dt^2}\frac{dx}{dt}+\frac{\partial^2 f}{\partial x\partial y}\left(\frac{d^2x}{dt^2}\frac{dy}{dt}+\frac{d^2y}{dt^2}\frac{dx}{dt}\right)+$$
$$+\frac{\partial^2 f}{\partial y^2}\frac{d^2y}{dt^2}\frac{dy}{dt}+\frac{\partial f}{\partial x}\frac{d^3x}{dt^3}+\frac{\partial f}{\partial y}\frac{d^3y}{dt^3},$$

and so on.

15.52. The formulae of 15.51 readily extend to functions of three or more variables. For instance

15.521. $$\frac{d}{dt}f(x,y,z) = \frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial z}\frac{dz}{dt}$$

and

15.522. $$\frac{d^2f}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}+\frac{dz}{dt}\frac{\partial}{\partial z}\right)^2 f+$$
$$+\frac{\partial f}{\partial x}\frac{d^2x}{dt^2}+\frac{\partial f}{\partial y}\frac{d^2y}{dt^2}+\frac{\partial f}{\partial z}\frac{d^2z}{dt^2}.$$

If x, y, z are functions, not of a single variable t but of, say, two variables u, v, then instead of 15.521 we have the pair of equations

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial u}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial u},$$
$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial v}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial v}.$$

15.53. In particular, if $u = x$, $v = y$, then

$$\left(\frac{\partial f}{\partial x}\right)_y = \frac{\partial f}{\partial x}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial x}, \qquad \left(\frac{\partial f}{\partial y}\right)_x = \frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial y},$$

where we have written $(\partial f/\partial x)_y$ to denote the x-derivative of f(x, y, z) when y is constant but z varies with x, and $\partial f/\partial x$ denotes as usual the x-derivative of f(x, y, z) when both y and z are constant, etc. The distinction is made in this way because if we sought to make it by writing df/dx for $(\partial f/\partial x)_y$ we should be liable to confuse the x-derivative of $f(x, y, z)$, when y is constant and z varies with x, with the x-derivative of $f(x, y, z)$, when both y and z vary with x. To make quite clear which of the variables are constant under differentiation it is sometimes necessary to place all such variables explicitly beside the sign of differentiation, e.g. we may denote the x-derivative of $f(x, y, z, u, v)$, when y, z are constant but u, v vary with x, by $(\partial f/\partial x)_{y,z}$. There are also other contexts in which it is desirable to make explicit what variables are constant during a differentiation. Consider, for instance, the relation between cartesian and polar coordinates

$$x = r\cos\theta, \qquad y = r\sin\theta.$$

If we regard x as a function of r and θ then $\dfrac{\partial x}{\partial\theta} = -r\sin\theta$, but if we regard x as a function of y and θ, i.e. $x = y\cot\theta$, then

$$\frac{\partial x}{\partial\theta} = -y\operatorname{cosec}^2\theta = -r/\sin\theta.$$

These derivatives are adequately distinguished by writing $(\partial x/\partial\theta)_r$ for the θ-derivative of x formed by keeping r constant, and $(\partial x/\partial\theta)_y$ for the θ-derivative of x keeping y constant; thus $\left(\dfrac{\partial x}{\partial\theta}\right)_r = -r\sin\theta$, $\left(\dfrac{\partial x}{\partial\theta}\right)_y = -r/\sin\theta$. Of course the subscript notation is unnecessary if all the variables of the function differentiated are made apparent. The distinction we made between $(\partial f/\partial x)_y$ and $\partial f/\partial x$ could be made by writing $\dfrac{\partial}{\partial x}f\{x, y, z(x, y)\}$ for the former and $\dfrac{\partial}{\partial x}f(x, y, z)$ for the latter, and similarly the distinction between $\left(\dfrac{\partial x}{\partial\theta}\right)_r$ and $\left(\dfrac{\partial x}{\partial\theta}\right)_y$ is equally well made by writing $\dfrac{\partial}{\partial\theta}x(\theta, r)$ and $\dfrac{\partial}{\partial\theta}x(\theta, y)$ for the θ-derivatives formed by keeping r constant and y constant respectively.

15.54. If u and v are differentiable functions of x and y, then

$$\frac{\partial u}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial u}=1, \qquad \frac{\partial v}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}\frac{\partial y}{\partial u}=0;$$

$$\frac{\partial v}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial v}{\partial y}\frac{\partial y}{\partial v}=1, \qquad \frac{\partial u}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial v}=0.$$

Let $u=f(x,\ y)$, $v=g(x,\ y)$, and consider these equations as expressing x and y as functions of u and v; then differentiating with respect to u we have

$$1=\frac{\partial f}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \quad \text{and} \quad 0=\frac{\partial g}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial g}{\partial y}\frac{\partial y}{\partial u},$$

and differentiating with respect to v,

$$0=\frac{\partial f}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial v} \quad \text{and} \quad 1 \quad \frac{\partial g}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial g}{\partial y}\frac{\partial y}{\partial u};$$

writing u for $f(x,y)$ and v for $g(x,y)$ we obtain the formulae stated. Of course $\partial x/\partial u$, $\partial x/\partial v$ here stand for the u- and v-derivatives of x regarded as a function of u and v whereas $\partial u/\partial x$, $\partial v/\partial x$ stand for the x-derivatives of u and v regarded as functions of x and y.

The technique by which these formulae were obtained is even more important than the formulae themselves.

15.55. When x and y are *linear* functions of t the formulae of § 15.51 take on a specially simple form. For if $x=a+ht$, $y=b+kt$ then $\dfrac{dx}{dt}=h$, $\dfrac{dy}{dt}=k$ and all higher derivatives of x and y are zero, so that

$$\begin{aligned} \frac{d}{dt}f(a+ht,\ b+kt) &= h\frac{\partial f}{\partial x}+k\frac{df}{\partial y}, \\ \frac{d^2f}{dt^2} &= \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^2 f, \\ \frac{d^3f}{dt^3} &= \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^3 f, \end{aligned}$$

and so on, where, on the right-hand side of each equation, x and y are replaced by $a+ht$ and $b+kt$ respectively, *after* differentiation.

Observe that if in such an expression as $\left(h\dfrac{\partial}{\partial x}+k\dfrac{\partial}{\partial y}\right)^n f(x,y)$ we replace x and y by $a+ht$ and $b+kt$ *after* differentiation and *then*

take $t = 0$, the result is just that of replacing x and y by a and b respectively *after* differentiation; but if for some r and s

$$\frac{\partial^{r+s}f(x, y)}{\partial x^r \partial y^s} = \phi(x, y)$$

then
$$\frac{\partial^{r+s}f(a, b)}{\partial a^r \partial b^s} = \phi(a, b),$$

and so the *result* of replacing x and y in $\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^n f(x, y)$ by a and b *after* differentiation is just $\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^n f(a, b)$. Hence

15.56.
$$\left[\frac{d^n}{dt^n}f(a+ht, b+kt)\right]_{t=0} = \left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^n f(a, b),$$

where $[d^nf/dt^n]_{t=0}$ denotes the value of d^nf/dt^n when t is replaced by zero *after* differentiation.

Formula 15.56 is in fact a particular case of the more general result, that if $x = a+ht$, $y = b+kt$ then

15.57
$$\frac{d^nf}{dt^n} = \left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^n f(a+ht, b+kt).$$

For

$$\frac{\partial}{\partial a}\phi(x, y) = \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial a}+\frac{\partial\phi}{\partial y}\frac{\partial y}{\partial a} = \frac{\partial\phi}{\partial x} \quad \text{since } \frac{\partial x}{\partial a} = 1,\ \frac{\partial y}{\partial a} = 0$$

and
$$\frac{\partial}{\partial b}\phi(x, y) = \frac{\partial\phi}{\partial x}\frac{\partial x}{\partial b}+\frac{\partial\phi}{\partial y}\frac{\partial y}{\partial b} = \frac{\partial\phi}{\partial y}$$

and therefore

$$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)f(a+ht, b+kt) = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)f(x, y),$$

and if for some p,

$$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^p f(a+ht, b+kt) = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^p f(x, y),$$

then

$$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^{p+1} f(a+ht, b+kt)$$
$$= \left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)\left[\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^p f(a+ht, b+kt)\right]$$
$$= \left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)\left[\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^p f(x, y)\right]$$
$$= \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)\left[\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^p f(x, y)\right] = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{p+1} f(x, y)$$

and so

$$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^n f(a+ht,\, b+kt) = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^n f(x,y)$$

for any n, which proves 15.57.

15.6. Taylor's theorem

If $f(x, y)$ is differentiable n times in the rectangle

$$(a, a+h)(b, b+k)$$

then there is a θ in $[0, 1]$ such that

$$f(a+h, b+k) = f(a, b)+\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)f(a, b)+$$
$$+\frac{1}{2!}\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^2 f(a, b)+\ldots+\frac{1}{(n-1)!}\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^{n-1} f(a, b)+$$
$$+\frac{1}{n!}\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}\right)^n f(a+\theta h, b+\theta k).$$

Let $\phi(t) = f(a+ht, b+kt)$, then $\phi(t)$ is differentiable n times and so by Maclaurin's theorem there is a θ in $[0, 1]$ such that

$$\phi(1) = \phi(0)+\phi'(0)+\frac{1}{2!}\phi''(0)+\ldots+\frac{1}{(n-1)!}\phi^{n-1}(0)+\frac{1}{n!}\phi^n(\theta),$$

whence by 15.56 and 15.57 (taking $t = \theta$) the result follows.

Exactly the same proof applies to a function of any number of variables.

15.7. A function of any number of variables $f(x, y, z,\ldots)$ is said to be homogeneous of order m if for any $x, y, z,\ldots$, and any t,

$$f(xt, yt, zt,\ldots) = t^m f(x, y, z,\ldots).$$

15.71. Euler's theorem

If $f(x, y, z,\ldots)$ is a homogeneous differentiable function of order m, then

$$\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+z\frac{\partial}{\partial z}+\ldots\right)^n f(x, y, z,\ldots)$$
$$= m(m-1)(m-2)\ldots(m-n+1)f(x, y, z,\ldots).$$

Let $u = xt$, $v = yt$, $w = zt,\ldots$ so that $u, v, w,\ldots$ are linear functions of t with $\frac{du}{dt} = x$, $\frac{dv}{dt} = y$, $\frac{dw}{dt} = z,\ldots$.

Since $f(x, y, z,...)$ is homogeneous,

$$f(u, v, w,...) = t^m f(x, y, z,...),$$

and so differentiating n times with respect to t,

$$\left(x\frac{\partial}{\partial u}+y\frac{\partial}{\partial v}+z\frac{\partial}{\partial w}+...\right)^n f(u, v, w,...)$$
$$= m(m-1)...(m-n+1)t^{m-n}f(x, y, z,...).$$

This is true for any t; taking $t = 1$ it follows from 15.55 (since $u, v, w,...$ take the values $x, y, z,...$ when $t = 1$) that

$$\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+z\frac{\partial}{\partial z}+...\right)^n f(x, y, z,...)$$
$$= m(m-1)...(m-n+1)f(x,y,z,...).$$

In particular, taking $n = 1$,

$$x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}+... = mf,$$

and, if m is a positive integer, taking $n = m+1$,

$$\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+...\right)^{m+1} f(x, y, z,...) = 0.$$

15.8. If $u(x, y)$ and $v(x, y)$ are differentiable functions then the determinant

$$\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix},$$

is called the *Jacobian* of the functions u, v. We denote the Jacobian of u, v by $\dfrac{\partial(u, v)}{\partial(x, y)}$ or $\partial(u, v)/\partial(x, y)$.

The definition readily extends to any number of functions, e.g. the Jacobian of $u(x, y, z)$, $v(x, y, z)$, $w(x, y, z)$ is

$$\partial(u, v, w)/\partial(x, y, z) \quad \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial z} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} & \dfrac{\partial v}{\partial z} \\ \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} & \dfrac{\partial w}{\partial z} \end{vmatrix},$$

As our notation has anticipated the Jacobian plays a role for sets

of functions very similar to that played by the derivative for a single function. The following theorems show how close this analogy is.

15.81. If $f_1(x, y)$, $f_2(x, y)$, $u(x, y)$, and $v(x, y)$ are differentiable functions then

$$\frac{\partial(f_1, f_2)}{\partial(x, y)} = \frac{\partial(f_1, f_2)}{\partial(u, v)} \times \frac{\partial(u, v)}{\partial(x, y)}.$$

For

$$\frac{\partial(f_1, f_2)}{\partial(u, v)} \times \frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_1}{\partial v} \\[2ex] \dfrac{\partial f_2}{\partial u} & \dfrac{\partial f_2}{\partial v} \end{vmatrix} \times \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{\partial f_1}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial f_1}{\partial v}\dfrac{\partial v}{\partial x} & \dfrac{\partial f_1}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial f_1}{\partial v}\dfrac{\partial v}{\partial y} \\[2ex] \dfrac{\partial f_2}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial f_2}{\partial v}\dfrac{\partial v}{\partial x} & \dfrac{\partial f_2}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial f_2}{\partial v}\dfrac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\[2ex] \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{vmatrix} = \frac{\partial(f_1, f_2)}{\partial(x, y)},$$

since

$$\frac{\partial f_r}{\partial x} = \frac{\partial f_r}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f_r}{\partial v}\frac{\partial v}{\partial x}, \qquad \frac{\partial f_r}{\partial y} = \frac{\partial f_r}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f_r}{\partial v}\frac{\partial v}{\partial y}, \qquad r = 1, 2.$$

The same argument establishes the analogous formula for any number of functions.

15.82. Inversion of a functional relation

If $y = f(x, u, v)$ and $\left|\dfrac{\partial f}{\partial x}\right| \geqslant \alpha > 0$ for all (x, u, v) in some interval $I = (x_0, x_1)(u_0, u_1)(v_0, v_1)$ then we can determine a function $x = \phi(y, u, v)$ such that $f\{\phi(y, u, v), u, v\} = y$ for all u, v, y.

By 15.24, $f_x(x, u, v)$ is of constant sign in I. Suppose that $\dfrac{\partial f}{\partial x} > 0$ in I; then $f(x, u, v)$ is steadily increasing with x and so if $x_0 < x < x_1$, $f(x, u, v)$ lies between $f(x_0, u, v)$ and $f(x_1, u, v)$. But $f(x, u, v)$ is continuous and therefore $f(x, u, v)$ takes each value between $f(x_0, u, v)$ and $f(x_1, u, v)$ once and once only; thus to any given

u, v in $(u_0, u_1)(v_0, v_1)$ and y between $f(x_0, u, v)$ and $f(x_1, u, v)$ there corresponds a unique x, which we denote by $\phi(y, u, v)$. By definition $\phi(y, u, v)$ satisfies $f\{\phi(y, u, v), u, v\} = y$ for all u, v, and y. *Furthermore, if $f(x, u, v)$ is differentiable then $\phi(y, u, v)$ is differentiable.*

For if $y = f(x, u, v)$ and $f(x+h, u+p, v+q) = y+k$, then by Taylor's theorem,

$$k = \left(h\frac{\partial}{\partial x}+p\frac{\partial}{\partial u}+q\frac{\partial}{\partial v}\right)f(x+\theta h, u+\theta p, v+\theta q).$$

Denoting $f_x(x+\theta h, u+\theta p, v+\theta q)$ for brevity by $f_x(\theta)$, and similarly for the u- and v-derivatives, we have

$$h = \{k-pf_u(\theta)-qf_v(\theta)\}/f_x(\theta). \tag{i}$$

Since f_x, f_u, f_v are bounded this shows that h is small when p, q, and k are small, and so x is a continuous function of y, u, v.

Taking $p = q = 0$ in (i) we have $h/k = 1/f_x(\theta) \to 1/f_x(x, u, v)$ as $k \to 0$, since when $k \to 0$, $h \to 0$ and so $x+\theta h \to x$. (Observe that the condition $|f_x| \geqslant \alpha > 0$ is required to justify this step; see 13.551, example iv). Similarly, taking $k = p = 0$,

$$h/q = -f_v(\theta)/f_x(\theta) \to -f_v/f_x,$$

and taking $k = q = 0$,

$$h/p = -f_u(\theta)/f_x(\theta) \to -f_u/f_x,$$

so that x is a differentiable function of y, u, v with y-, u-, and v-derivatives $1/f_x$, $-f_u/f_x$, and $-f_v/f_x$ respectively.

Observe that once we *know* that x is a differentiable function of y, u, v then the values of the derivatives can be obtained from the formulae 15.53.

For $f(x, u, v)-y$ is *constantly* zero for all u, v, y and $x = \phi(u, v, y)$ and so

$$\frac{\partial}{\partial u}\{f(x, u, v)-y\} = \frac{\partial}{\partial v}\{f(x, u, v)-y\} = \frac{\partial}{\partial y}\{f(x, u, v)-y\} = 0,$$

giving

$$\frac{\partial f}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial f}{\partial u} = 0, \qquad \frac{\partial f}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial f}{\partial v} = 0, \qquad \frac{\partial f}{\partial x}\frac{\partial x}{\partial y}-1 = 0,$$

whence

$$\frac{\partial x}{\partial u} = -\frac{\partial f}{\partial u}\Big/\frac{\partial f}{\partial x}, \qquad \frac{\partial x}{\partial v} = -\frac{\partial f}{\partial v}\Big/\frac{\partial f}{\partial x}, \qquad \frac{\partial x}{\partial y} = 1\Big/\frac{\partial f}{\partial x}$$

as above.

15.83. If $x = f(u, v, t)$ and $y = g(u, v, t)$ are differentiable functions in some interval R and if

$$\frac{\partial(f, g)}{\partial(u, v)} \geqslant \alpha > 0 \quad \text{in } R,$$

then we can determine differentiable functions $u = F(x, y, t)$, $v = G(x, y, t)$ such that

$$f\{F(x, y, t), G(x, y, t), t\} = x \quad \text{and} \quad g\{F(x, y, t), G(x, y, t), t\} = y$$

for all x, y, t, such that (u, v, t) lies in R.

Suppose that, at a *given* point u, v, t,

$\phi(u, v, t)$ is the greatest of the four numbers $\left|\frac{\partial f}{\partial u}\right.$ $\left|\frac{\partial g}{\partial u}\right.$

and let $\alpha = 8\beta^2$, $\beta > 0$, then

$$2\phi^2 \geqslant \left|\frac{\partial f}{\partial u}\frac{\partial g}{\partial v} - \frac{\partial f}{\partial v}\frac{\partial g}{\partial u}\right. \geqslant$$

and so $\phi \geqslant 2\beta$. Thus at *each* point of R *one* at least of the functions $|\partial f/\partial u|$, $|\partial f/\partial v|$, $|\partial g/\partial u|$, $|\partial g/\partial v|$ exceeds 2β, and therefore by 15.25 we can divide R into a finite number of sub-intervals in each of which one of the functions (at least) exceeds β throughout the sub-interval. Let ρ be one of the sub-intervals and suppose that $|\partial f/\partial u|$ is the function which exceeds β throughout ρ. Then by 15.82 we can determine, in ρ, a differentiable function $u = \phi(v, x, t)$ such that $x = f\{\phi(v, x, t), v, t\}$ for all v, t, x. We prove next that the equation $y = g\{\phi(v, x, t), v, t\}$ can be solved for v, determining v as a function of x, y, t, and hence from $u = \phi(v, x, t)$, u is determined as a function of x, y, t. Let $g\{\phi(v, x, t), v, t\} = \psi(v, x, t)$, then

$$\frac{\partial \psi}{\partial v} = \frac{\partial g}{\partial u}\frac{\partial \phi}{\partial v} + \frac{\partial g}{\partial v}; \tag{i}$$

but $f(u, v, t) - x$ is constantly zero for all v, t, x and $u = \phi(v, x, t)$, and so, differentiating with respect to v,

$$0 = \frac{\partial f}{\partial u}\frac{\partial \phi}{\partial v} + \frac{\partial f}{\partial v}. \tag{ii}$$

Multiplying (i) by $\partial f/\partial u$ and (ii) by $\partial g/\partial u$, and subtracting, we have

$$\frac{\partial f}{\partial u}\frac{\partial \psi}{\partial v} = \frac{\partial(f, g)}{\partial(u, v)} \quad \text{and so} \quad \left|\frac{\partial f}{\partial u}\right|\left|\frac{\partial \psi}{\partial v}\right| \geq \alpha.$$

But $\partial f/\partial u$ is continuous, and so $|\partial f/\partial u|$ is bounded, by M say, and therefore $|\partial \psi/\partial v| \geqslant \alpha/M$, whence, by 15.82, v is determined as a function of x, y, t.

Since u and v are functions of x, y, t in each sub-interval ρ, therefore u and v are functions of x, y, t throughout R. Denote these functions by $u = F(x, y, t)$, $v = G(x, y, t)$; F and G satisfy

$$f\{F(x, y, t), G(x, y, t), t\} = x,$$
$$g\{F(x, y, t), G(x, y, t), t\} = y$$

for all x, y, t, by definition.

It remains to show that $u = F(x, y, t)$ and $v = G(x, y, t)$ are differentiable.

Since $f(u, v, t)$ and $g(u, v, t)$ are differentiable, if

$$x+h = f(u+p, v+q, t+\tau) \quad \text{and} \quad y+k = g(u+p, v+q, t+\tau),$$

then by Taylor's theorem

$$h = \left(p\frac{\partial}{\partial u}+q\frac{\partial}{\partial v}+\tau\frac{\partial}{\partial t}\right)f(u+\theta_1 p, v+\theta_1 q, t+\theta_1 \tau)$$
$$= \left(p\frac{\partial}{\partial u}+q\frac{\partial}{\partial v}+\tau\frac{\partial}{\partial t}\right)f(\theta_1) \quad \text{(say)}$$

and

$$k = \left(p\frac{\partial}{\partial u}+q\frac{\partial}{\partial v}+\tau\frac{\partial}{\partial t}\right)g(u+\theta_2 p, v+\theta_2 q, t+\theta_2 \tau)$$
$$= \left(p\frac{\partial}{\partial u}+q\frac{\partial}{\partial v}+\tau\frac{\partial}{\partial t}\right)g(\theta_2) \quad \text{(say)}.$$

Multiplying the first equation by $\dfrac{\partial g(\theta_2)}{\partial v}$ and the second by $\dfrac{\partial f(\theta_1)}{\partial v}$, and subtracting, it follows that

$$h\frac{\partial g(\theta_2)}{\partial v} - k\frac{\partial f(\theta_1)}{\partial v} = p\left\{\frac{\partial f(\theta_1)}{\partial u}\frac{\partial g(\theta_2)}{\partial v} - \frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial u}\right\} +$$
$$+\tau\left\{\frac{\partial f(\theta_1)}{\partial t}\frac{\partial g(\theta_2)}{\partial v} - \frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial t}\right\}. \quad \text{(iii)}$$

Since

$$\frac{\partial f(\theta_1)}{\partial u}\frac{\partial g(\theta_2)}{\partial v} - \frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial u} \to \frac{\partial f}{\partial u}\frac{\partial g}{\partial v} - \frac{\partial f}{\partial v}\frac{\partial g}{\partial u} \quad \text{as} \quad p, q, \tau \to 0$$

$\left|\dfrac{\partial(f,g)}{\partial(u,v)}\right| \geqslant \alpha$, therefore when p, q, τ are sufficiently small,

$$\left|\frac{\partial f(\theta_1)}{\partial u}\frac{\partial g(\theta_2)}{\partial v}-\frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial u}\right| \geqslant \tfrac{1}{2}\alpha:$$

furthermore $\partial f/\partial u$, $\partial f/\partial v$, $\partial g/\partial u$, $\partial g/\partial v$ being continuous, are bounded, and so equation (iii) shows that when $h, k, \tau \to 0$ then necessarily $p \to 0$, which proves that u is a *continuous* function of x, y, t. Similarly v is a continuous function of x, y, t.

Taking in turn $k = \tau = 0$, $h = \tau = 0$, and $h = k = 0$ in equation (iii) we have

$$p/h = \frac{\partial g(\theta_2)}{\partial v}\Big/\left\{\frac{\partial f(\theta_1)}{\partial u}\frac{\partial g(\theta_2)}{\partial v}-\frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial u}\right\} \to \frac{\partial g}{\partial v}\Big/\frac{\partial(f,g)}{\partial(u,v)} \quad \text{as} \quad h \to 0$$

(since $h \to 0$ entails $p \to 0$ and $q \to 0$ so that $u+\theta_1 p \to u, v+\theta_1 q \to v$, etc.), and

$$p/k = -\frac{\partial f(\theta_1)}{\partial v}\Big/\left\{\frac{\partial f(\theta_1)}{\partial u}\frac{\partial g(\theta_2)}{\partial v}-\frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial u}\right\} \to -\frac{\partial f}{\partial v}\Big/\frac{\partial(f,g)}{\partial(u,v)} \quad \text{as} \quad k \to 0,$$

and

$$p/\tau = -\left\{\frac{\partial f(\theta_1)}{\partial t}\frac{\partial g(\theta_2)}{\partial v}-\frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial t}\right\}\Big/\left\{\frac{\partial f(\theta_1)}{\partial u}\frac{\partial g(\theta_2)}{\partial v}-\frac{\partial f(\theta_1)}{\partial v}\frac{\partial g(\theta_2)}{\partial u}\right\}$$

$$\to -\frac{\partial(f,g)}{\partial(t,v)}\Big/\frac{\partial(f,g)}{\partial(u,v)} \quad \text{as} \quad \tau \to 0$$

which proves that u is a differentiable function of x, y, t.

Similarly v is a differentiable function of x, y, t.

Once we have established that u and v are differentiable functions the derivatives are given by 15.53.

For instance, since $f(u, v, t)-x$ and $g(u, v, t)-y$ are constantly zero for all x, t and $u = F(x, y, t)$, $v = G(x, y, t)$, therefore

$$\frac{\partial f}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = 1, \qquad \frac{\partial g}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial g}{\partial v}\frac{\partial v}{\partial x} = 0,$$

whence
$$\frac{\partial u}{\partial x} = \frac{\partial g}{\partial v}\Big/\frac{\partial(f,g)}{\partial(u,v)}$$

as above, and so on.

15.84. If $x = f(u, v, w, t)$, $y = g(u, v, w, t)$, $z = h(u, v, w, t)$ are differentiable functions in some interval R and if

$$\left|\frac{\partial(f,g,h)}{\partial(u,v,w)}\right| \geqslant \alpha > 0 \quad \text{in } R,$$

then we can determine differentiable functions $u = F(x, y, t)$, $v = G(x, y, t)$, $w = H(x, y, t)$ satisfying $x = f(F, G, H, t)$, $y = g(F, G, H, t)$, $z = h(F, G, H, t)$ for all values of x, y, t such that (u, v, w, t) lies in R.

Let J_u, J_v, J_w be the co-factors of $\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w}$ in the determinant $\frac{\partial(f, g, h)}{\partial(u, v, w)} = J$, so that

$$J = \frac{\partial f}{\partial u}J_u + \frac{\partial f}{\partial v}J_v + \frac{\partial f}{\partial w}J_w.$$

Let M be a bound of all the derivatives $\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \ldots, \frac{\partial h}{\partial w}$, then since $|J| \geqslant \alpha > 0$, at *each* point of R one of the numbers $|J_u|$, $|J_v|$, $|J_w|$ exceeds α/M, and so by 15.25 we can divide R into a finite number of sub-intervals in each of which at least one of $|J_u|$, $|J_v|$, $|J_w|$ exceeds $\alpha/2M$ throughout the sub-interval. Let ρ be one of the sub-intervals, and suppose that $|J_u| \geqslant \alpha/2M$ in ρ. Then by 15.83 we can determine v and w as functions $\phi(u, y, t)$, $\psi(u, z, t)$; inserting these values of v and w in $x = f(u, v, w, t)$ we obtain an equation of the form $x = X(u, y, z, t)$. We show next that this equation can be solved for u, whence u, v, w are determined as functions of x, y, z, t.

We have
$$\frac{\partial X}{\partial u} = \frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\frac{\partial \phi}{\partial u} + \frac{\partial f}{\partial w}\frac{\partial \psi}{\partial u},$$
$$0 = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}\frac{\partial \phi}{\partial u} + \frac{\partial g}{\partial w}\frac{\partial \psi}{\partial u},$$
$$0 = \frac{\partial h}{\partial u} + \frac{\partial h}{\partial v}\frac{\partial \phi}{\partial u} + \frac{\partial h}{\partial w}\frac{\partial \psi}{\partial u},$$
whence
$$\frac{\partial(f, g, h)}{\partial(u, v, w)} = \frac{\partial X}{\partial u}\frac{\partial(g, h)}{\partial(v, w)} = \frac{\partial X}{\partial u}J_u.$$

But $\alpha/2M \leqslant |J_u| < 2M^2$ and so $|\partial X/\partial u| \geqslant \alpha/2M^3$, from which it follows that the equation $x = X(u, y, z, t)$ can be solved for u, and therefore u, v, w are determined as functions of x, y, z, t; the proof that these functions are differentiable proceeds exactly as in 15.83. Furthermore, Theorem 15.84 can be extended step by step to any number of functions, the proof in each case following that of 15.84 exactly.

15.85. The solution of an equation in the neighbourhood of a given point

If $f(x, u, v)$ is a differentiable function in R, $f(a, l, m) = 0$, where the point (a, l, m) lies in R and $|f_x(a, l, m)| > 0$, then we can determine a differentiable function $x = \phi(u, v)$ satisfying $f\{\phi(u, v), u, v\} = 0$ at all points (u, v) *near* (l, m) and such that $\phi(l, m) = a$.

Let $f_x(a, l, m) = 2\delta$, and suppose that $\delta > 0$. Since $f_x(x, u, v)$ is continuous we can determine an interval ρ such that at all points of ρ, $|f_x(x, u, v) - f_x(a, l, m)| < \delta$, and therefore $f_x(x, u, v) > \delta$. Accordingly $f(x, u, v)$ is increasing in x and so, if $x < a < X$, then

$$f(x, l, m) < f(a, l, m) < f(X, l, m);$$

but $f(a, l, m) = 0$ and so

$$f(x, l, m) < 0 < f(X, l, m).$$

Since $f(x, u, v)$ is continuous and $f(x, l, m) < 0$, $f(X, l, m) > 0$, then, for any u, v sufficiently close to (l, m), $f(x, u, v) < 0$ and $f(X, u, v) > 0$. Let σ be an interval in which these inequalities hold. Then for each given pair (u, v) in σ, the continuous steadily increasing function $f(x, u, v)$ vanishes for one and only one value of x; denote this value of x by $\phi(u, v)$. Then $f\{\phi(u, v), u, v\} = 0$ for any u, v in σ, and, moreover, since there is only one value of x for which $f(x, l, m) = 0$, and $f(a, l, m) = 0$, therefore $\phi(l, m) = a$.

Since $f(x, u, v)$ is differentiable, if $f(a+h, l+p, m+q) = 0$ then by Taylor's theorem

$$\begin{aligned} 0 &= f(a+h, l+p, m+q) - f(a, l, m) \\ &= \left(h\frac{\partial}{\partial x} + p\frac{\partial}{\partial u} + q\frac{\partial}{\partial v}\right) f(a+\theta h, l+\theta p, m+\theta q) \\ &= \left(h\frac{\partial}{\partial x} + p\frac{\partial}{\partial u} + q\frac{\partial}{\partial v}\right) f(\theta), \quad \text{say,} \end{aligned}$$

i.e.
$$h f_x(\theta) + p f_u(\theta) + q f_v(\theta) = 0,$$

whence the fact that x is a differentiable function follows as in 15.82.

15.86. If any number of functions

$$f(x, y, z, \ldots, u, v, w, \ldots), \qquad g(x, y, z, \ldots, u, v, w, \ldots),$$
$$h(x, y, z, \ldots, u, v, w, \ldots), \quad \ldots$$

are all differentiable in R, and if at the point $(a,b,c,\dots,l,m,n,\dots)$ in R

$$f(a,b,c,\dots,l,m,n,\dots) = g(a,b,c,\dots,l,m,n,\dots)$$
$$= h(a,b,c,\dots,l,m,n) = \dots = 0$$

and
$$\frac{\partial(f,g,h,\dots)}{\partial(x,y,z,\dots)} \neq 0,$$

then we can determine differentiable functions

$$x = F(u,v,w,\dots), \quad y = G(u,v,w,\dots), \quad z = H(u,v,w,\dots), \quad \dots$$

such that

$$a = F(l,m,n,\dots), \quad b = G(l,m,n,\dots), \quad c = H(l,m,n,\dots), \quad \dots$$

and

$$f(F,G,H,\dots,u,v,w,\dots) = g(F,G,H,\dots,u,v,w,\dots)$$
$$= h(F,G,H,\dots,u,v,w,\dots) = \dots = 0$$

at all points $u,v,w,\dots$ near $l,m,n,\dots$.

The theorem is proved if we can show that its truth for $n+1$ functions follows from that for n functions. For simplicity in notation we shall illustrate this step by proving the extension from two to three functions.

Let J_x, J_y, J_z be the co-factors of $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$, $\dfrac{\partial f}{\partial z}$ in the determinant $\dfrac{\partial(f,g,h)}{\partial(x,y,z)} = J$. Then $J = f_x J_x + f_y J_y + f_z J_z$; since J is non-zero at (a,b,c,l,m,n), it follows that at this point at least one of the co-factors is non-zero. Suppose that this cofactor is J_x. Then $\dfrac{\partial(g,h)}{\partial(y,z)} \neq 0$ and so (assuming the truth of 15.86 for the case of two functions) we can determine functions $y = \phi(x,u,v,w)$, $z = \psi(x,u,v,w)$ which take the values b and c at the point (a,l,m,n) and which satisfy $g(x,\phi,\psi,u,v,w) = h(x,\phi,\psi,u,v,w) = 0$ at all points near (a,l,m,n). It remains to determine x as a function of u, v, w which will satisfy

$$f(x,\phi,\psi,u,v,w) = 0 \quad \text{at all points } u, v, w \text{ near } (l,m,n).$$

Write
$$f(x,\phi,\psi,u,v,w) = X(x,u,v,w).$$

Then
$$X(a,l,m,n) = f(a,b,c,l,m,n) = 0$$

and
$$\frac{\partial X}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial \phi}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial \psi}{\partial x};$$

but
$$0 = \frac{\partial g}{\partial x}+\frac{\partial g}{\partial y}\frac{\partial \phi}{\partial x}+\frac{\partial g}{\partial z}\frac{\partial \psi}{\partial x}$$
and
$$0 = \frac{\partial h}{\partial x}+\frac{\partial h}{\partial y}\frac{\partial \phi}{\partial x}+\frac{\partial h}{\partial z}\frac{\partial \psi}{\partial x},$$
$$\therefore \quad \frac{\partial X}{\partial x}\frac{\partial(g,h)}{\partial(y,z)} = \frac{\partial(f,g,h)}{\partial(x,y,z)}.$$

Since neither $\dfrac{\partial(g,h)}{\partial(y,z)}$ nor $\dfrac{\partial(f,g,h)}{\partial(x,y,z)}$ vanishes at (a, b, c, l, m, n), therefore $\partial X/\partial x$ is non-zero at this point, and so by 15.85 we can determine a function $x = F(u, v, w)$ which takes the value a at the point (l, m, n) and which satisfies

$$X(F, u, v, w) = 0 \quad \text{at all points } (u, v, w) \text{ near } (l, m, n).$$

Thus we have determined x, y, z as functions of u, v, w in the neighbourhood of (l, m, n) and if we denote these functions by F, G, H respectively, then by definition

$$f(F,G,H,u,v,w) = g(F,G,H,u,v,w) = h(F,G,H,u,v,w) = 0$$

for all u, v, w near l, m, n and F, G, H take the values a, b, c at the point (l, m, n).

The proof that F, G, and H are differentiable proceeds exactly as in Theorem 15.83.

15.87. If $u = u(x, y, z)$, $v = v(x, y, z)$, $w = w(x, y, z)$ are differentiable functions in an interval R and if

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0$$

for all points of R, then there is a functional relation between u, v, and w which does not contain x, y, or z.

(This theorem is true for any number of functions.)

Proof. The following possibilities must be distinguished:

(α) all the Jacobians $\dfrac{\partial(v,w)}{\partial(y,z)}$, $\dfrac{\partial(u,w)}{\partial(x,z)}$, $\dfrac{\partial(u,v)}{\partial(x,y)}$ vanish at all points of R;

(β) at least one of these Jacobians is non-zero somewhere in R.

(α) may be further divided into two cases:

(α_1) all the derivatives $\dfrac{\partial u}{\partial x}, \dfrac{\partial u}{\partial y}, \ldots, \dfrac{\partial v}{\partial z}, \ldots, \dfrac{\partial w}{\partial z}$ vanish throughout R;

(α_2) at least one of these derivatives is non-zero somewhere in R.

Consider (β) first. Suppose that $\left|\frac{\partial(v, w)}{\partial(y, z)}\right| - 2\delta > 0$ at a point (a, b, c) in R; then because $\frac{\partial(v, w)}{\partial(y, z)}$ is continuous there is some interval ρ containing (a, b, c) throughout which $\left|\frac{\partial(v, w)}{\partial(y, z)}\right| \geqslant \delta$. Hence by 15.83 we can determine y and z as functions $\phi(x, v, w)$ and $\psi(x, v, w)$ satisfying $v(x, \phi, \psi) = v$ and $w(x, \phi, \psi) = w$ for all x, v, w corresponding to points in ρ. We shall show that the functional relation $u = u(x, \phi, \psi)$ is independent of x and is therefore a relation between u, v, w alone; this is ensured if the x-derivative of $u(x, \phi, \psi)$ is zero, i.e. if $\lambda = 0$, where

$$\lambda = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial \phi}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial \psi}{\partial x}.$$

Now
$$\frac{\partial v}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial \phi}{\partial x} + \frac{\partial v}{\partial z}\frac{\partial \psi}{\partial x} = 0$$

and
$$\frac{\partial w}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial \phi}{\partial x} + \frac{\partial w}{\partial z}\frac{\partial \psi}{\partial x} = 0,$$

and therefore
$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \lambda\frac{\partial(v, w)}{\partial(y, z)}.$$

Since $\frac{\partial(u, v, w)}{\partial(x, y, z)} = 0$ throughout ρ and $\left|\frac{\partial(v, w)}{\partial(y, z)}\right| \geqslant \delta$ throughout ρ, therefore $\lambda = 0$ throughout ρ and there is a functional relation between u, v, w for all values of u, v, w corresponding to points in ρ.

In case (α_1), u, v, and w are mere constants and so we can establish innumerable functional relations between them.

There remains to consider case (α_2).

Let $\partial u/\partial x$ be the derivative which has a non-zero value somewhere in R; then since $\partial u/\partial x$ is continuous there is some interval σ throughout which $|\partial u/\partial x| \geqslant \alpha > 0$. Hence by 15.82 we can determine x as a function $\chi(y, z, u)$ satisfying $u = u(\chi, y, z)$.

We shall show next that the relations $v = v(\chi, y, z) = V(y, z, u)$, say, and $w = w(\chi, y, z) = W(y, z, u)$ are both independent of y and z so that u, v, and w are connected by the relations

$$v = V(u), \qquad w = W(u).$$

For
$$\frac{\partial V}{\partial y} = \frac{\partial v}{\partial x}\frac{\partial \chi}{\partial y} + \frac{\partial v}{\partial y} \quad \text{and} \quad 0 = \frac{\partial u}{\partial x}\frac{\partial \chi}{\partial y} + \frac{\partial u}{\partial y}$$

$$\frac{\partial V}{\partial y}\frac{\partial u}{\partial x} = \frac{\partial(u, v)}{\partial(x, y)} = 0,$$

$\left|\frac{\partial u}{\partial x}\right| \geqslant \alpha$, we have $\frac{\partial V}{\partial y} = 0$ throughout σ. Similarly $\frac{\partial V}{\partial x} = \frac{\partial W}{\partial x} = \frac{\partial W}{\partial y} = 0$ and so V and W are independent of y and z, which completes the proof.

Example. Let a function $t(x)$ be defined by the integral

$$t(x) = \int\limits_0^x \frac{1}{1+u^2}\,du, \quad \text{so that} \quad t(0) = 0, \quad t'(x) = \frac{1}{1+x^2}$$

then $$t(x)+t(y) = t[(x+y)/(1-xy)], \quad xy < 1.$$

For if $u = t(x)+t(y)$, $v = t[(x+y)/(1-xy)]$ then v is differentiable in the region $xy < 1$ or the region $xy > 1$ and

$$\frac{\partial u}{\partial x} = t'(x) = \frac{1}{1+x^2}$$

$$\frac{\partial v}{\partial x} = \frac{1}{1+[(x+y)^2/(1-xy)^2]}\left\{\frac{1}{1-xy}+\frac{(x+y)y}{(1-xy)^2}\right\}$$

$$= \frac{1+y^2}{(1+x^2)(1+y^2)} = \frac{1}{1+x^2}, \quad \text{provided } xy \neq 1$$

and

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{1}{1+x^2} & \frac{1}{1+y^2} \\ \frac{1}{1+x^2} & \frac{1}{1+y^2} \end{vmatrix} \quad \text{for all } x \text{ and } y,\ xy \neq 1$$

so that by 15.86 there is a functional relation between u, v, say $u = \phi(v)$, in *either* of the regions $xy < 1$, $xy > 1$, and so

$$\phi[t\{(x+y)/(1-xy)\}] = t(x)+t(y) \quad \text{for all } x,\ y \text{ provided } xy \neq 1.$$

Take $y = 0$, which lies in the region $xy < 1$, then $t(x) = \phi(t(x))$ for all x, and therefore

$$\phi[t\{(x+y)/(1-xy)\}] = t\{(x+y)/(1-xy)\}, \quad xy < 1,$$

whence $$t(x)+t(y) = t\{(x+y)/(1-xy)\}, \quad xy < 1.$$

15.88. Singular points

By Taylor's theorem a repeatedly differentiable function $f(x, y)$ may be expanded in the form

$$f(X,Y) = f(x,y)+\{(X-x)f_x+(Y-y)f_y\}+$$
$$+\frac{1}{2!}\{(X-x)^2f_{xx}+2(X-x)(Y-y)f_{xy}+(Y-y)^2f_{yy}\}+...+R_n$$
$$= f(x,y)+L_1(X,Y)+\frac{1}{2!}L_2(X,Y)+...+R_n, \quad \text{say}$$

where

$$R_n = \frac{1}{n!}\left\{(X-x)\frac{\partial}{\partial x}+(Y-y)\frac{\partial}{\partial y}\right\}^n f\{x(1-\theta)+X\theta,\ y(1-\theta)+Y\theta$$

If $f(x, y) = 0$ then the curve $f(X, Y) = 0$ passes through the point (x, y), and $(X-x)f_x+(Y-y)f_y = 0$ is the equation of the tangent at (x, y), provided f_x, f_y are not both zero, since on the curve $f(X, Y) = 0$ we have $f_X+f_Y\dfrac{dY}{dX} = 0$, so that at the point (x, y) we have $\dfrac{dY}{dX} = -f_x/f_y$ when $f_y \neq 0$, and $\dfrac{dX}{dY} = -f_y/f_x$ when $f_x \neq 0$.

If $f(x, y)$, f_x, f_y are all zero, then (x, y) is said to be a *singular point* on the curve, the type of singularity depending upon the terms of second degree in $X-x$, $Y-y$.

If $L_2(X, Y)$ has a pair of factors $(X-x)p_r+(Y-y)q_r$, $r = 1, 2$, the curve is said to have two *branches* at the singular point (x, y) and both the lines $(X-x)p_r+(Y-y)q_r = 0$ are tangents to $f(X, Y) = 0$ at the point (x, y). For differentiating with respect to X the identity $f_X+f_Y(dY/dX) = 0$ we have

$$f_{XX}+2f_{XY}\frac{dY}{dX}+f_{YY}\left(\frac{dY}{dX}\right)^2+f_Y\frac{d^2Y}{dX^2} = 0,$$

and taking $X = x$, $Y = y$ this gives the equation

$$f_{xx}+2f_{xy}\frac{dY}{dX}+f_{yy}\left(\frac{dY}{dX}\right)^2 = 0,$$

which by hypothesis has the solutions

$$p_r+q_r\frac{dY}{dX} = 0, \qquad r = 1, 2,$$

showing that both the lines

$$(X-x)p_r+(Y-y)q_r = 0, \quad r = 1, 2,$$

are tangents at (x, y).

If $L_2 = 0$ has no linear factors then no point near (x, y), save (x, y) itself, lies on the curve. For writing $X-x = h$, $Y-y = k$ we have

$$f(X, Y) = f(x+h, y+k)$$
$$= \frac{1}{2!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^2 f(x, y)+\frac{1}{3!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^3 f(x+\theta h, y+\theta k).$$

At points near (x, y), but different from (x, y), one of h, k at least is not zero; suppose $k \neq 0$, and write r, α for the polar coordinates of the point (h, k). Then

$$f(X, Y)$$
$$= \frac{r^2}{2!}\left\{\left(\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}\right)^2 f+\frac{r}{3}\left(\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}\right)^3 f(x+\theta h, y+\theta k)\right\}$$

and so $$\lim_{r\to 0}\frac{2f(X, Y)}{r^2} = \left(\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}\right)^2 f,$$

and since L_2 has no factors, $\left(\cos\alpha\frac{\partial}{\partial x}+\sin\alpha\frac{\partial}{\partial y}\right)^2 f$ is not zero for any value of α, and therefore $f(X, Y)$ is not zero for all sufficiently small values of r. Accordingly (x, y) is an *isolated* point on the curve.

If L_2 is a perfect square then the curve may have two branches at (x, y), with coincident tangents, or (x, y) may be an isolated point on the curve. A point where two tangents coincide is called a *cusp*. The two cases are illustrated in Examples (i) and (iv) below.

If all the second derivatives f_{xx}, f_{xy}, and f_{yy} are also zero the type of singularity is determined by the terms of the third degree. In this case we may have one or three tangents (of which two or more may coincide) but (x, y) cannot now be an isolated point, unless all the third degree terms are absent, since a cubic polynomial has necessarily one linear factor.

Examples. (i) The semi-cubical parabola $y^3-x^2 = 0$ has a singular point at the origin; the singularity is a cusp, the y-axis touching the curve at the origin.

(ii) The folium of Descartes, $x^3+y^3-3xy = 0$, has a singular point at the origin, the curve having two branches at the origin, the axes of coordinates both being tangents.

(iii) The curve $x^3+y^3+x^2+y^2 = 0$ has an isolated singularity at the origin.

(iv) The curve $y^2+x^4-x^3y^2 = 0$ has an isolated singularity at the origin; for writing $y = tx^2$ we have $x^4(t^2+1-t^2x) = 0$, and so, either $x = 0$, or $x = 1+1/t^2 \geqslant 1$, which shows that x can take no value < 1, except 0. Moreover, if $y = 0$ then $x = 0$, and therefore the origin is an isolated point.

15.9. If $f(x, y)$ is x-continuous in (a, b), then for any given value of y, $f(x, y)$ is a continuous function of a single variable x and therefore integrable in (a, b); the value of the integral, of course, depends upon y.

15.91. If f(x, y) is differentiable in (a, b)(c, d) then the function

$$\mathbf{F(x, y) = \int_a^x f(t, y)\,dt}$$

is differentiable in (a, b)(c, d) and

$$\mathbf{F_x(x, y) = f(x, y), \qquad F_y(x, y) = \int_a^x f_y(t, y)\,dt.}$$

Since $f(x, y)$ is differentiable it is x-continuous and so $F(x, y)$ is determined for any (x, y) in $(a, b)(c, d)$.

Let M be a bound of $f(x, y)$ in $(a, b)(c, d)$; then

$$F(X, Y)-F(x, y) = \int_a^X f(t, Y)\,dt - \int^x f(t, y)\,dt$$

$$= \int_a^x \{f(t, Y)-f(t, y)\}\,dt + \int_x^X f(t, Y)\,dt.$$

Since $f(x, y)$ is continuous $f(t, Y)-f(t, y) = 0(p)$ for any (t, y), (t, Y) in $(a, b)(c, d)$ such that $Y-y = 0(q)$; hence if $X-x = 0(p)$, $Y-y = 0(q)$ then

$$F(X, Y)-F(x, y) = (x-a)0(p)+M\,0(p) = \{(b-a)+M\}0(p)$$

so that $F(x, y)$ is continuous.

Furthermore

$$\frac{F(X, y)-F(x, y)}{X-x} = \frac{1}{X-x}\int_x^X f(t, y)\,dt$$

$$= f(c, y)$$

by the *mean-value theorem* for integrals, where c lies between x and

X (and here depends also upon y). Since $f(x, y)$ is continuous, $f(c, y)-f(x, y) = 0(p)$ for any (x, y), (X, y) in $(a, b)(c, d)$ such that $X-x = 0(q)$, and therefore

$$\frac{F(X, y)-F(x, y)}{X-x}-f(x, y) = 0(p),$$

which proves that $F(x, y)$ is x-differentiable with x-derivative $f(x, y)$.

Finally,

$$F(x, Y)-F(x, y) = \int_a^x f(t, Y)\,dt - \int_a^x f(t, y)\,dt$$

$$= \int_a^x \{f(t, Y)-f(t, y)\}\,dt$$

and so
$$\frac{F(x, Y)-F(x, y)}{Y-y} = \int_a^x \frac{f(t, Y)-f(t, y)}{Y-y}\,dt.$$

But $f(x, y)$ is differentiable, and so

$$\frac{f(t, Y)-f(t, y)}{Y-y} = f_y(t, y)+0(p)$$

for all (t, y), (t, Y) in $(a, b)(c, d)$ such that $Y-y = 0(q)$.

Hence, if $Y-y = 0(q)$,

$$\frac{F(x, Y)-F(x, y)}{Y-y} - \int_a^x f_y(t, y)\,dt = \int_a^x 0(p)\,dt = (x-a)0(p)$$
$$= (b-a)0(p)$$

and so $F(x, y)$ is also y-differentiable with derivative $\int_a^x f_y(t, y)\,dt$.

In particular if
$$F(y) = \int_a^b f(t, y)\,dt$$

then
$$F'(y) = \int_a^b f_y(t, y)\,dt.$$

15.92. If $f(x, y)$ is differentiable in $(a, b)(c, d)$, y_1, y_2 are differentiable functions of y whose values lie in (a, b) when y lies in (c, d), and if

$$\phi(y) = \int_{y_1}^{y_2} f(t, y)\,dt,$$

then

$$\varphi'(y) = \int_{y_1}^{y_2} f_y(t, y)\,dt + \frac{dy_2}{dy} f(y_2, y) - \frac{dy_1}{dy} f(y_1, y).$$

Let $F(x, y) = \int_a^x f(t, y)\,dt$, then $\phi(y) = F(y_2, y) - F(y_1, y)$. But $F_x(x, y) = f(x, y)$, $F_y(x, y) = \int_a^x f_y(t, y)\,dt$, and therefore

$$\begin{aligned}\phi'(y) &= F_x(y_2, y)\frac{dy_2}{dy} + F_y(y_2, y) - F_x(y_1, y)\frac{dy_1}{dy} - F_y(y_1, y)\\ &= f(y_2, y)\frac{dy_2}{dy} - f(y_1, y)\frac{dy_1}{dy} + \int_a^{y_2} f_y(t, y)\,dt - \int_a^{y_1} f_y(t, y)\,dt\\ &= f(y_2, y)\frac{dy_2}{dy} - f(y_1, y)\frac{dy_1}{dy} + \int_{y_1}^{y_2} f_y(t, y)\,dt.\end{aligned}$$

XVI

MAXIMA AND MINIMA

FREE AND RESTRICTED MAXIMUM AND MINIMUM VALUES OF FUNCTIONS OF SEVERAL VARIABLES. MAXIMIN AND MINIMAX VALUES. ENVELOPES OF FAMILIES OF PLANE CURVES

16. The definition of a maximum or minimum value of a function of several variables is the same as that for a function of a single variable. The function $f(x, y, z, \ldots)$ is said to have a maximum at the point $(a, b, c, \ldots)$, or $f(a, b, c, \ldots)$ is said to be a maximum value of $f(x, y, z, \ldots)$, if

$$f(x, y, z, \ldots) < f(a, b, c, \ldots)$$

for all points $(x, y, z, \ldots)$ sufficiently near to $(a, b, c, \ldots)$.

Similarly, $f(x, y, z, \ldots)$ has a minimum value at $(a, b, c, \ldots)$ if

$$f(x, y, z, \ldots) > f(a, b, c, \ldots)$$

for all points $(x, y, z, \ldots)$ sufficiently near to $(a, b, c, \ldots)$.

Since the analysis is the same whatever the number of variables, we shall confine our attention to functions of not more than three variables.

16.1. If $f(x, y, z)$ has a maximum at (a, b, c) and if $x(t), y(t), z(t)$ is any continuous curve passing through (a, b, c), so that $x(t) = a$, $y(t) = b$, and $z(t) = c$ when $t = t_0$, then the function

$$\phi(t) = f\{x(t), y(t), z(t)\}$$

has a maximum at the point t_0. Conversely, if $\phi(t)$ has a maximum at t_0 for every continuous curve $x(t), y(t), z(t)$ which passes through (a, b, c) then $f(x, y, z)$ has a maximum at (a, b, c).

For if t is near to t_0 then $x(t), y(t), z(t)$ are near to $x(t_0), y(t_0), z(t_0)$ respectively and therefore

$$\phi(t) = f\{x(t), y(t), z(t)\} < f\{x(t_0), y(t_0), z(t_0)\} = \phi(t_0).$$

16.11. Throughout 16.1 we may replace 'maximum' by 'minimum', making the appropriate change in the inequality.

16.2. The function $f(x, y, z)$ has a maximum at (a, b, c) if

$$f_a(a, b, c) = f_b(a, b, c) = f_c(a, b, c) = 0$$

and $$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}+l\frac{\partial}{\partial c}\right)^2 f(a, b, c) < 0$$

for any h, k, l (not all simultaneously zero).

Let $x(t)$, $y(t)$, $z(t)$ be any curve passing through (a, b, c), the functions $x(t)$, $y(t)$, $z(t)$ taking the values a, b, c at the point t_0. We suppose that $f(x, y, z)$ and $x(t)$, $y(t)$, $z(t)$ are differentiable twice. Write $f\{x(t), y(t), z(t)\} = \phi(t)$, then $\phi(t)$ has a maximum at t_0 if, at this point, $df/dt = 0$ and $d^2f/dt^2 < 0$; but

$$\frac{df}{dt} = f_x\frac{dx}{dt}+f_y\frac{dy}{dt}+f_z\frac{dz}{dt}$$

and

$$\frac{d^2f}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}+\frac{dz}{dt}\frac{\partial}{\partial z}\right)^2 f(x, y, z)+f_x\frac{d^2x}{dt^2}+f_y\frac{d^2y}{dt^2}+f_z\frac{d^2z}{dt^2},$$

and when $t = t_0$, x, y, and z take the values a, b, c. Therefore the condition $df/dt = 0$ at t_0 is equivalent to

$$f_a\frac{dx}{dt}+f_b\frac{dy}{dt}+f_c\frac{dz}{dt} = 0.$$

This equation must be true for *any* curve $x(t)$, $y(t)$, $z(t)$ passing through a, b, c, and so for *any* values of dx/dt, dy/dt, dz/dt (for if dx/dt, dy/dt, dz/dt have the values h, k, l at t_0 then we may take $x(t) = a+h(t-t_0)$, $y(t) = b+k(t-t_0)$, and $z(t) = c+l(t-t_0)$); taking first $dx/dt = 1$, $dy/dt = 0$, $dz/dt = 0$ we find $f_a = 0$, and similarly $f_b = f_c = 0$. Hence at the point t_0

$$\frac{d^2f}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial a}+\frac{dy}{dt}\frac{\partial}{\partial b}+\frac{dz}{dt}\frac{\partial}{\partial c}\right)^2 f(a, b, c),$$

and since $d^2f/dt^2 < 0$ for all curves through (a, b, c) we have, writing h, k, l for dx/dt, dy/dt, dz/dt,

$$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}+l\frac{\partial}{\partial c}\right)^2 f(a, b, c) < 0$$

for any h, k, l (not all simultaneously zero).

16.21. In the same way we can show that $f(x, y, z)$ is a minimum at (a, b, c) if

$$f_a = f_b = f_c = 0$$

and $$\left(h\frac{\partial}{\partial a}+k\frac{\partial}{\partial b}+l\frac{\partial}{\partial c}\right)^2 f(a, b, c) > 0$$

for all h, k, l (not all simultaneously zero).

EXAMPLE. To find the maximum and minimum values of the function
$$\phi(x, y) = ax^2+2hxy+by^2+2gx+2fy+c.$$
We have
$$\phi_x = 2(ax+hy+g) = 0,$$
$$\phi_y = 2(hx+by+f) = 0$$
so that
$$\frac{x}{hf-bg} = \frac{y}{gh-af} = \frac{-1}{ab-h^2}. \tag{i}$$
Furthermore, $\phi_{xx} = 2a$, $\phi_{xy} = 2h$, $\phi_{yy} = 2b$, and so, writing x' for dx/dt, etc.,
$$\frac{d^2\phi}{dt^2} = 2(ax'^2+2hx'y'+by'^2).$$

If $a = b = 0$, then there is neither a maximum nor minimum value since $x'y'$ is positive when x', y' have the same sign, and negative otherwise.

If $a \neq 0$, then
$$\frac{d^2\phi}{dt^2} = \frac{2}{a}\{(ax'+hy')^2+(ab-h^2)y'^2\}.$$

If $ab-h^2 < 0$ then $d^2\phi/dt^2$ has opposite signs when $y' = 0$ and when $ax'+hy' = 0$; if $ab-h^2 = 0$ then $d^2\phi/dt^2$ vanishes when $ax'+hy' = 0$. Hence ϕ has neither a maximum nor minimum value when $ab-h^2 \leqslant 0$.

If $ab-h^2 > 0$ then $d^2\phi/dt^2$ has the same sign as a, vanishing only when x', y' vanish together, and so, when $a > 0$, $\phi(x, y)$ has a minimum value, and when $a < 0$, $\phi(x, y)$ has a maximum value, at the point (x, y) given by equations (i).

Observe that when $ab-h^2 > 0$ then a and b have the same sign (and neither is zero).

16.3. *The maximum and minimum values of a function when some of the variables are subject to certain restrictions.*

It will suffice to consider a function $f(x, y, u, v)$, where u and v are given in terms of x, y by the functional relations
$$F_1(x, y, u, v) = 0, \qquad F_2(x, y, u, v) = 0.$$
If the equations $F_1 = 0$, $F_2 = 0$ were solved for u and v in terms of x and y we could replace u and v by their values in terms of x and y in the function $f(x, y, u, v)$, and if $\phi(x, y)$ is the function so determined then the problem of determining the maximum and

minimum values of $f(x, y, u, v)$ is reduced to that, considered in 16.2, of determining the maximum and minimum values of $\phi(x, y)$ where x and y are independent. It may however be impossible or highly impracticable to solve the equations $F_1 = 0$, $F_2 = 0$, in which case the maximum and minimum values of $f(x, y, u, v)$ are determined as follows.

16.31. The function $f(x, y, u, v)$ has a maximum at $x = a$, $y = b$ if for any curve $x(t)$, $y(t)$ passing through (a, b)

$$\frac{df}{dt} = 0 \quad \text{and} \quad \frac{d^2f}{dt^2} < 0.$$

Now
$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial u}\frac{du}{dt}+\frac{\partial f}{\partial v}\frac{dv}{dt}, \tag{i}$$

where dx/dt, dy/dt are arbitrary, but du/dt, dv/dt satisfy

$$0 = \frac{dF_1}{dt} = \frac{\partial F_1}{\partial x}\frac{dx}{dt}+\frac{\partial F_1}{\partial y}\frac{dy}{dt}+\frac{\partial F_1}{\partial u}\frac{du}{dt}+\frac{\partial F_1}{\partial v}\frac{dv}{dt} \tag{ii}$$

and
$$0 = \frac{dF_2}{dt} = \frac{\partial F_2}{\partial x}\frac{dx}{dt}+\frac{\partial F_2}{\partial y}\frac{dy}{dt}+\frac{\partial F_2}{\partial u}\frac{du}{dt}+\frac{\partial F_2}{\partial v}\frac{dv}{dt}. \tag{iii}$$

The elimination of du/dt, dv/dt from equations (ii), (iii) and $df/dt = 0$ is most simply effected by introducing two new variables λ_1, λ_2 and adding the equation $df/dt = 0$ to the sum of λ_1 times equation (ii) and λ_2 times equation (iii), giving

$$\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}+\frac{du}{dt}\frac{\partial}{\partial u}+\frac{dv}{dt}\frac{\partial}{\partial v}\right)(f+\lambda_1 F_1+\lambda_2 F_2) = 0; \tag{iv}$$

Let the variables λ_1, λ_2 satisfy the two linear conditions

$$\frac{\partial f}{\partial u}+\lambda_1\frac{\partial F_1}{\partial u}+\lambda_2\frac{\partial F_2}{\partial u} = 0, \tag{γ}$$

$$\frac{\partial f}{\partial v}+\lambda_1\frac{\partial F_1}{\partial v}+\lambda_2\frac{\partial F_2}{\partial v} = 0. \tag{δ}$$

From (iv) and (γ), (δ) it follows that

$$\left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}\right)(f+\lambda_1 F_1+\lambda_2 F_2) = 0. \tag{v}$$

Since dx/dt, dy/dt may take any values we please, let $\frac{dx}{dt} = 1$, $\frac{dy}{dt} = 0$ and $\frac{dx}{dt} = 0$, $\frac{dy}{dt} = 1$ in turn and we find

$$\frac{\partial f}{\partial x}+\lambda_1\frac{\partial F_1}{\partial x}+\lambda_2\frac{\partial F_2}{\partial x} = 0, \tag{α}$$

$$\frac{\partial f}{\partial y}+\lambda_1\frac{\partial F_1}{\partial y}+\lambda_2\frac{\partial F_2}{\partial y} = 0. \tag{β}$$

The four equations (α), (β), (γ), (δ) express the condition $\frac{df}{dt} = 0$. *Observe that if* $\Phi(x, y, u, v) = f+\lambda_1 F_1+\lambda_2 F_2$ *then the same equations* (α), (β), (γ), (δ) *express the condition* $\frac{d\Phi}{dt} = 0$ *when* x, y, u, *and* v *are all independent variables.*

Consider next the second derivative of $f(x, y, u, v)$. We have

$$\frac{d^2f}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}+\frac{du}{dt}\frac{\partial}{\partial u}+\frac{dv}{dt}\frac{\partial}{\partial v}\right)^2 f+$$
$$+\frac{\partial f}{\partial x}\frac{d^2x}{dt^2}+\frac{\partial f}{\partial y}\frac{d^2y}{dt^2}+\frac{\partial f}{\partial u}\frac{d^2u}{dt^2}+\frac{\partial f}{\partial v}\frac{d^2v}{dt^2}$$

and, for $r = 1, 2$,

$$0 = \frac{d^2F_r}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}+\frac{du}{dt}\frac{\partial}{\partial u}+\frac{dv}{dt}\frac{\partial}{\partial v}\right)^2 F_r+$$
$$+\frac{\partial F_r}{\partial x}\frac{d^2x}{dt^2}+\frac{\partial F_r}{\partial y}\frac{d^2y}{dt^2}+\frac{\partial F_r}{\partial u}\frac{d^2u}{dt^2}+\frac{\partial F_r}{\partial v}\frac{d^2v}{dt^2},$$

whence adding the first equation to λ_1 times the second and λ_2 times the third equation, and utilizing equations (α), (β), (γ), and (δ),

$$\frac{d^2f}{dt^2} = \left(\frac{dx}{dt}\frac{\partial}{\partial x}+\frac{dy}{dt}\frac{\partial}{\partial y}+\frac{du}{dt}\frac{\partial}{\partial u}+\frac{dv}{dt}\frac{\partial}{\partial v}\right)^2 (f+\lambda_1 F_1+\lambda_2 F_2).$$

Thus if Φ denotes the function $f+\lambda_1 F_1+\lambda_2 F_2$ then the conditions that $f(x, y, u, v)$ has a maximum at (a, b), subject to the restrictions $F_1(x, y, u, v) = F_2(x, y, u, v) = 0$, are just that the function $\Phi(x, y, u, v)$ has a maximum at $x = a$, $y = b$, without restriction on the variables.

The values of λ_1 and λ_2 in the function Φ are determined by the equations (α) and (β). To determine whether or not $d^2\Phi/dt^2$ is of

constant sign it is generally necessary to eliminate du/dt and dv/dt by means of equations (ii) and (iii) above.

EXAMPLES. 1. To determine the maximum and minimum values of x^2+y^2 subject to the condition

$$Ax^2+2Bxy+Cy^2-H = 0, \quad AH > 0.$$

Write $\Phi(x, y) = x^2+y^2+\lambda(Ax^2+2Bxy+Cy^2-H)$

and $f(x, y) = x^2+y^2.$

Then $\dfrac{\partial\Phi}{\partial x} = \dfrac{\partial\Phi}{\partial y} = 0$ if

$$x+\lambda(Ax+By) = 0, \tag{i}$$

$$y+\lambda(Bx+Cy) = 0. \tag{ii}$$

Multiplying the first equation by x, the second by y, and adding, we find

$$f(x, y)+\lambda H = 0. \tag{iii}$$

Since $H \neq 0$, x and y cannot be zero simultaneously, and so $\lambda \neq 0$, and equating the value of x/y from (i) and (ii) we have

$$B^2\lambda^2 = (1+\lambda A)(1+\lambda C). \tag{iv}$$

Furthermore, along the curve $x(t)$, $y(t)$,

$$\frac{d^2\Phi}{dt^2} = 2\left\{(1+\lambda A)\left(\frac{dx}{dt}\right)^2+2\lambda B\frac{dx}{dt}\frac{dy}{dt}+(1+\lambda C)\left(\frac{dy}{dt}\right)^2\right\}$$

$$= 2(1+\lambda A)\left\{\frac{dx}{dt}+\frac{\lambda B}{1+\lambda A}\frac{dy}{dt}\right\}^2 \quad \text{using (iv).}$$

But $x(t)$, $y(t)$ satisfy $Ax^2+2Bxy+Cy^2 = H$, and so, denoting dx/dt, dy/dt by x', y' respectively we have

$$x'(Ax+By)+y'(Bx+Cy) = 0,$$

whence from (i), (ii) $\quad xx'+yy' = 0.$

Hence by (i) again, $\quad (1+\lambda A)y' = \lambda Bx'$

(since x, y are not both zero) and so

$$x'+\{\lambda B/(1+\lambda A)\}y' = x'\{1+\lambda^2B^2/(1+\lambda A)^2\}.$$

Therefore $\quad \dfrac{d^2\Phi}{dt^2} = \dfrac{2}{(1+\lambda A)}\{(1+\lambda A)^2+\lambda^2B^2\}^2x'^2.$

Hence $d^2\Phi/dt^2$ is of constant sign, the sign depending upon that of $1+\lambda A$.

Write $1+\lambda A = \mu$ then $\lambda = (\mu-1)/A$ and so by (iv) μ satisfies

$$B^2(\mu-1)^2 = A\mu\{A+C(\mu-1)\},$$

i.e. $$Q(\mu) \equiv (AC-B^2)\mu^2+\mu(A^2-AC+2B^2)-B^2 = 0.$$

The discriminant of the quadratic equation is

$$(A^2-AC+2B^2)^2+4B^2(AC-B^2) = A^2(A-C)^2+4A^2B^2$$

and so the roots are real. The product of the roots is $B^2/(B^2-AC)$ (provided $B^2 \neq AC$), and so if $B^2 < AC$ the roots are of opposite sign, and less than unity since $Q(0) < 0$, $Q(1) > 0$, and if $B^2 > AC$ both the roots are positive but only one makes $\lambda A < 0$, and so $\lambda H < 0$. If $B^2 = AC$ then the equation for μ is linear and μ has the positive value $B^2/(A^2+B^2)$.

Thus if $B^2 < AC$ (so that $Ax^2+2Bxy+Cy^2-H = 0$ is the equation of an ellipse) then x^2+y^2 has one maximum and one minimum value; if $B^2 > AC$ (so that the equation represents a hyperbola) then x^2+y^2 has one minimum value; and if $B^2 = AC$ (when the equation represents a pair of lines) then x^2+y^2 has one minimum value. These values of $f(x, y) = x^2+y^2$ are given by

$$B^2f^2 = (H-Af)(H-Cf),$$

obtained by eliminating λ between (iii) and (iv).

2. To find the maximum and minimum values of $x^2+y^2+z^2$ subject to the conditions $x+y+z = 1$, $xyz+1 = 0$.

Write $\Phi(x, y, z) = x^2+y^2+z^2+\lambda_1(x+y+z-1)+\lambda_2(xyz+1)$.

Then $\dfrac{\partial\Phi}{\partial x} = \dfrac{\partial\Phi}{\partial y} = \dfrac{\partial\Phi}{\partial z} = 0$ if

$$2x+\lambda_1+\lambda_2 yz = 0, \tag{i}$$

$$2y+\lambda_1+\lambda_2 zx = 0, \tag{ii}$$

$$2z+\lambda_1+\lambda_2 xy = 0. \tag{iii}$$

Hence $2(x-y)-\lambda_2 z(x-y) = 0$ and so, *either* $x = y$ *or* $\lambda_2 z = 2$.

If $\lambda_2 = 2/z$ then $\lambda_1 = -2(x+y)$ and so, from the third equation,

$$2z-2(x+y)+2xy/z = 0.$$

But $x+y = 1-z$ and $xy = -1/z$, and therefore $2z-1-1/z^2 = 0$, whence $2z^3-z^2-1 = 0$, i.e. $z^2(z-1)+z^3-1 = 0$, and so

$$(z-1)\{2z^2+z+1\} = 0.$$

Since $$2z^2+z+1 = \tfrac{1}{2}(2z+\tfrac{1}{2})^2+\tfrac{7}{8} > 0,$$

therefore $z = 1$, from which it follows that $x+y = 0$ and $xy = -1$, whence $x = \pm 1$, $y = \mp 1$, and $\lambda_1 = 0$, $\lambda_2 = 2$.

On the other hand, if $x = y$, then $z = 1-2x$, and $z = -1/x^2$, which gives $2x^3-x^2-1 = 0$ and so $x = 1$; hence $y = 1$ and $z = -1$ and therefore from (i) and (iii), $2+\lambda_1-\lambda_2 = 0$ and $-2+\lambda_1+\lambda_2 = 0$, whence $\lambda_1 = 0$, $\lambda_2 = 2$. Thus the only solutions are

$$\lambda_1 = 0, \quad \lambda_2 = 2; \qquad x = 1, \quad y = 1, \quad z = -1;$$
$$x = 1, \quad y = -1, \quad z = 1; \qquad x = -1, \quad y = 1, \quad z = 1.$$

Now

$$\frac{d^2\Phi}{dt^2} = 2\left(\frac{dx}{dt}\right)^2+2\left(\frac{dy}{dt}\right)^2+2\left(\frac{dz}{dt}\right)^2+2\lambda_2 x\frac{dy}{dt}\frac{dz}{dt}+$$
$$+2\lambda_2 y\frac{dz}{dt}\frac{dx}{dt}+2\lambda_2 z\frac{dx}{dt}\frac{dy}{dt}.$$

Hence when $x = 1$, $y = 1$, $z = -1$

$$\frac{d^2\Phi}{dt^2} = 2\left\{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2+2\frac{dy}{dt}\frac{dz}{dt}+2\frac{dz}{dt}\frac{dx}{dt}-2\frac{dx}{dt}\frac{dy}{dt}\right\}$$
$$= 8\left(\frac{dx}{dt}\right)^2 > 0;$$

for we have

$$\frac{dx}{dt}+\frac{dy}{dt}+\frac{dz}{dt} = 0, \qquad yz\frac{dx}{dt}+zx\frac{dy}{dt}+xy\frac{dz}{dt} = 0$$

and so at $x = 1$. $y = 1$, $z = -1$

$$\frac{dx}{dt}+\frac{dy}{dt} = \frac{dz}{dt},$$

whence $$\frac{dz}{dt} = 0, \quad \text{and} \quad \frac{dy}{dt} = -\frac{dx}{dt}.$$

Similarly, when $x = 1$, $y = -1$, $z = 1$ and $x = -1$, $y = 1$, $z = 1$,

$$\frac{d^2\Phi}{dt^2} > 0.$$

Thus $x^2+y^2+z^2$ has three minimum values, all equal.

16.4. The conditions for a maximum or minimum value of a function $f(x, y)$ at the point (a, b) are that $f_a(a, b) = f_b(a, b) = 0$ and that

$$f_{aa}h^2+2f_{ab}hk+f_{bb}k^2$$

be of constant sign for all values of h, k (not both zero).

Since

$$f_{aa}h^2+2f_{ab}hk+f_{bb}k^2 = \{(f_{aa}h+f_{ab}k)^2+(f_{aa}f_{bb}-f_{ab}^2)k^2\}/f_{aa}$$

this expression is of constant sign only if $f_{aa}f_{bb}-f_{ab}^2 > 0$; for if $f_{aa}f_{bb}-f_{ab}^2 < 0$ then the values corresponding to $k = 0$ and $f_{aa}h+f_{ab}k = 0$ are of opposite sign, and if $f_{aa}f_{bb}-f_{ab}^2 = 0$ then $f_{aa}h^2+2f_{ab}hk+f_{bb}k^2$ is zero for all values of h and k satisfying $f_{aa}h+f_{ab}k = 0$.

Thus the conditions for a maximum or minimum value of $f(x, y)$ at the point (a, b) may be expressed in the form

$$f_a = f_b = 0, \qquad f_{aa}f_{bb}-f_{ab}^2 > 0;$$

when these conditions are satisfied then $f(x, y)$ has a minimum at (a, b) if f_{aa} is positive, and a maximum if f_{aa} is negative. (Note that f_{aa} and f_{bb} necessarily have the same sign when $f_{aa}f_{bb} > f_{ab}^2$.)

If $f_a = f_b = 0$ and $f_{aa}f_{bb}-f_{ab}^2 < 0$, then $f(x, y)$ is said to have a *maximin* value at (a, b) if *either* $f_{aa} > 0$ *or* $f_{bb} > 0$ and a *minimax* value at (a, b) if *either* $f_{aa} < 0$ *or* $f_{bb} < 0$. Since f_{aa} and f_{bb} may have opposite signs it may happen that $f(x, y)$ is both maximin and minimax at the one point (a, b).

16.41. A maximin value of a function is a maximum of restricted minimum values of the function, and a minimax is a minimum of restricted maximum values.

Let $f(a, b)$ be a maximin value of $f(x, y)$; then either f_{aa} or f_{bb} exceeds zero, and we shall suppose it is the former.

Since $f_x(x, y) = 0$ when $x = a$, $y = b$ and $f_{xx}(x, y) > 0$ at this point then, in the neighbourhood of (a, b) there is a differentiable function $\phi(y)$ such that $\phi(b) = a$ and $f_x\{\phi(y), y\} = 0$ for all y near b.

Since $f_{xx}\{\phi(b), b\} > 0$ it follows that the function $f(x, y)$, regarded as a function of x, has a minimum value for $x = \phi(y)$, for *all* values of y, sufficiently near b. We show next that these minimum values $f\{\phi(y), y\}$ have a maximum value at $y = b$.

Since $f(a, b)$ is a maximin value of $f(x, y)$ therefore

$$f_x(x, y) = f_y(x, y) = 0, \quad \text{when} \quad x = a, y = b.$$

Hence $D_y f\{\phi(y), y\} = f_x\phi'+f_y = 0$ when $y = b$, and since $f_x\{\phi(y), y\} = 0$ for all y near b, therefore $f_{xx}\phi'+f_{xy} = 0$, and so

$$\begin{aligned} D_y^2 f\{\phi(y), y\} &= f_{yy}+2f_{xy}\phi'+f_{xx}\phi'^2+f_x\phi'' \\ &= \{f_{yy}f_{xx}-f_{xy}^2\}/f_{xx} < 0, \quad \text{when} \quad y = b, \end{aligned}$$

which proves that $f\{\phi(y), y\}$ is maximum when $y = b$.

Similarly, if $f(a, b)$ is a minimax value of $f(x, y)$, with $f_{aa} < 0$, then we can find a function $\psi(y)$ so that $f(x, y)$, regarded as a function of x, is maximum for $x = \psi(y)$, for all values of y near b, and $f\{\psi(y), y\}$ is minimum when $y = b$.

Example. The function $x^2+4xy-y^2$ has both a maximin and a minimax value at the origin; for if $f(x, y) = x^2+4xy-y^2$ then $f_x = 2x+4y$, $f_y = 4x-2y$, both of which vanish at the origin, and $f_{xx} = 2$, $f_{xy} = 4$, $f_{yy} = -2$, so that $f_{xx}f_{yy}-f_{xy}^2 < 0$. Thus $x^2+4xy-y^2$ is both maximin and minimax at the origin.

Observe that, as a function of x, $x^2+4xy-y^2$, is minimum for $x = -2y$, for all values of y and $(-2y)^2+4y(-2y)-y^2$ is maximum at the origin. Similarly, as a function of y, $x^2+4xy-y^2$ is maximum for $y = 2x$ and $x^2+4x(2x)-(2x)^2$ is minimum at the origin.

16.5. Envelopes

For each value of a the equation $f(x, y, a) = 0$ determines a plane curve C_a; accordingly $f(x, y, a) = 0$ is called the equation of a *family* of curves, or, specifically, a *one-parameter family* of curves.

If for each value of a there is a point $x(a)$, $y(a)$ on the curve C_a such that the curve Γ whose parametric equations are $x = x(a)$, $y = y(a)$ touches C_a at the point $x(a)$, $y(a)$, then the family of curves $f(x, y, a) = 0$ is said to admit an envelope, *and the curve Γ is called an* envelope *of the family.*

16.51. To find whether a family of curves $f(x, y, a) = 0$ admits an envelope, and to determine the envelope when it exists.

Keeping a constant, the slope of the tangent at a point (x, y) on the curve of parameter a is dy/dx, given by

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx} = 0; \tag{i}$$

let $\phi(a)$, $\psi(a)$ be the point of the curve where it is touched by the envelope, then as a varies the equation of the envelope is $x = \phi(a)$, $y = \psi(a)$ and the tangent to the envelope at the point a is of slope

$$\frac{dy}{dx} = \frac{dy}{da}\Big/\frac{dx}{da} = \psi'(a)/\phi'(a). \tag{ii}$$

But for each value of a the point $\phi(a)$, $\psi(a)$ lies on the curve $f(x, y, a) = 0$ and therefore $f\{\phi(a), \psi(a), a\} = 0$, whence

$$\frac{\partial f}{\partial x}\phi'(a)+\frac{\partial f}{\partial y}\psi'(a)+\frac{\partial f}{\partial a} = 0. \tag{iii}$$

Since the curve and the envelope have the same tangent at their point of contact, we have from (i) and (ii)

$$\frac{\partial f}{\partial x}\phi'(a)+\frac{\partial f}{\partial y}\psi'(a) = 0$$

and therefore from (iii) $\qquad \dfrac{\partial f}{\partial a} = 0.$

Thus any point (x, y) on the envelope satisfies the pair of equations

$$\mathbf{f(x, y, a) = 0, \qquad \frac{\partial}{\partial a}f(x, y, a) = 0;}$$

if these equations can be solved for x and y, determining x and y as functions of a, $\phi^*(a)$ and $\psi^*(a)$, say, then the curve $x = \phi^*(a)$, $y = \psi^*(a)$ *may* be the envelope. The point $x = \phi^*(a)$, $y = \psi^*(a)$ is called a *characteristic* point of the curve $f(x, y, a) = 0$. There may also be other curves whose points lie on $f(x, y, a) = 0$, $\dfrac{\partial}{\partial a}f(x, y, a) = 0$; for if, corresponding to any a, there is a point $x = \lambda(a)$, $y = \mu(a)$ on $f(x, y, a) = 0$ such that

$$\frac{\partial}{\partial x}f(x, y, a) = \frac{\partial}{\partial y}f(x, y, a) = 0$$

then at such a point (by (iii)) $\partial f/\partial a = 0$, whatever values $\lambda'(a)$, $\mu'(a)$ may have. A point of $f(x, y, a) = 0$ where $\partial f/\partial x = \partial f/\partial y = 0$ is called a *singular* point of the curve; thus we have shown that

not only points of the envelope, but also singular points of curves of the family, satisfy the equations

$$f(x, y, a) = 0, \qquad \frac{\partial}{\partial a}f(x, y, a) = 0.$$

At a singular point the tangent is not determinate. For if

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0,$$

then the equation for dy/dx, viz.

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx}=0$$

is satisfied by any value of dy/dx.

16.52. A point of intersection, (x', y'), *if any*, of the curves $f(x, y, a) = 0$ and $f(x, y, a') = 0$ is given by the pair of equations $f(x', y', a) = 0, f(x', y', a') = 0$. But, for a certain c between a and a', $f(x', y', a') = f(x', y', a)+(a'-a)\frac{\partial}{\partial a}f(x', y', c)$, and so x', y' satisfy $f(x', y', a) = 0, \frac{\partial}{\partial a}f(x', y', c) = 0$. Hence as $a' \to a$, x' and y' tend to x and y, satisfying

$$f(x, y, a) = 0, \qquad \frac{\partial}{\partial a}f(x, y, a) = 0$$

(for $c \to a$ when $a' \to a$), i.e. *as $a' \to a$ a point of intersection of $f(x, y, a) = 0$ and $f(x, y, a') = 0$, when these curves meet, tends to that point on $f(x, y, a) = 0$ where it has contact with the envelope.* The intersection of $f(x, y, a') = 0$ with $f(x, y, a) = 0$ as a' tends to a, cannot be used to *define* the characteristic points since the curves may not meet however small $|a'-a|$ may be.

EXAMPLES. 1. The envelope of the family of lines

$$a^2x-ay+c = 0, \quad a \neq 0$$

(where c is constant), is given by the equations

$$a^2x-ay+c = 0, \qquad 2ax-y = 0;$$

eliminating a between these equations, we have $y^2 = 4cx$, which is a true envelope, for its slope at the point c/a^2, $2c/a$ is

$$\frac{dy}{dx}=\frac{4c}{4c/a}=a,$$

which is also the slope of $a^2x-ay+c = 0$. The condition for a singular point is $a^2 = a = 0$ and so the family has no singular point.

Regarding $a^2x-ay+c = 0$ as a quadratic in a, we see that in general two lines (or none) pass through any given point (x, y), but if x and y satisfy $y^2 = 4cx$, the two lines coincide, and therefore just off the envelope $y^2 = 4cx$, the two lines are close together.

The point of intersection of two lines of the family $a^2x-ay+c=0$, $b^2x-by+c=0$ is $x'=c/ab$, $y'=(a+b)c/ab$, and as $b \to a$ then $x' \to c/a^2$, $y' \to 2c/a$, a point on the envelope.

2. The family $(y-a)^2 = x^3$ has no envelope. For if there were an envelope it would satisfy

$$(y-a)^2 = x^3, \qquad 2(y-a) = 0;$$

but these equations are satisfied only by points on the line $x = 0$. The derivatives of $(y-a)^2-x^3$ with respect to x and y are $-3x^2$ and $2(y-a)$, and these are zero for $x = 0$, $y = a$. Thus the point $x = 0$, $y = a$ on $(y-a)^2 = x^3$ is a singular point and the line $x = 0$ contains only singular points and is not an envelope.

3. The envelope of the family of parabolas $y = a(x-a)^2$ is given by

$$y = a(x-a)^2, \qquad (x-a)^2-2a(x-a) = 0,$$

i.e. by $x = a$, $y = 0$ or $x = 3a$, $y = 4a^3$. None of the curves has a singular point since the derivative with respect to y has the constant value unity. Thus both $y = 0$ and $x = 3a$, $y = 4a^3$, i.e. $y = \frac{4}{27}x^3$, are envelopes, so the envelope is made up of the cubic $y = \frac{4}{27}x^3$ and the line $y = 0$ which also belongs to the family itself. The enveloping cubic touches $y = a(x-a)^2$ at the point $x = 3a$, $y = 4a^3$ and meets it again at the point $x = \frac{3}{4}a$, $y = \frac{1}{16}a^3$, showing that an envelope may intersect the members of a family not only at the points of contact but at further points as well.

Through a general point (x, y) pass three parabolas (or one) of the family, their parameters given by the cubic

$$a^3-2a^2x+ax^2-y = 0,$$

but through points on the envelope $y = \frac{4}{27}x^3$ two of the parabolas coincide, showing that near the envelope two of the parabolas are close together; for when $y = \frac{4}{27}x^3$ the cubic for the parameters becomes $a^3-2a^2x+ax^2-\frac{4}{27}x^3 = 0$ which has the solutions $a = \frac{1}{3}x$ twice, and $a = \frac{4}{3}x$, so that through a point (x', y') near the envelope pass two parabolas both with parameters near $\frac{1}{3}x'$, if $|y'| < \frac{4}{27}|x'|^3$.

The parabolas $y = a(x-a)^2$, $y = b(x-b)^2$, when a and b have the same sign, meet at the points

$$x' = a+b+\sqrt{ab}, \qquad y' = ab(a+b+2\sqrt{ab})$$

and

$$x'' = a+b-\sqrt{ab}, \qquad y'' = ab(a+b-2\sqrt{ab}),$$

and when $b \to a$, $x' \to 3a$, $y' \to 4a^3$ and $x'' \to a$, $y'' \to 0$ (or vice

versa according as a is positive or negative) and both of these points lie on the envelope. If we impose the restriction $a \neq 0$ then the origin is no longer part of the envelope since this point on the line $y = 0$ is not then a point of contact with the family.

4. Show that the envelope of the line $x\sin t - y\cos t + f'(t) = 0$ is the evolute of the envelope of the line $x\cos t + y\sin t - f(t) = 0$, and that the arc length of the latter envelope is $|f'(t) + \int f(t)\,dt|$, provided $f(t)+f''(t)$ is of constant sign.

Proof. The envelope of the line

$$L(t) = x\cos t + y\sin t - f(t) = 0$$

is the locus of points which lie both on $L(t) = 0$ and on

$$L'(t) = -x\sin t + y\cos t - f'(t) = 0.$$

The lines $L(t) = 0$, $L'(t) = 0$ are perpendicular, and so, since the envelope of $L(t) = 0$ is tangential to $L(t) = 0$, therefore $L'(t) = 0$ is normal to the envelope of $L(t) = 0$, and so the envelope of $L'(t) = 0$ is the locus of the centres of curvature of the envelope of $L(t) = 0$ (see § 10.96).

Hence, since the envelope of $L'(t) = 0$ is the locus of the point of intersection of the lines $L'(t) = 0$, $L''(t) = 0$, it follows that the radius of curvature of the envelope of $L(t) = 0$ is the distance between the point of intersection of $L(t) = 0$, $L'(t) = 0$ and the point of intersection of $L'(t) = 0$, $L''(t) = 0$. Denoting these points by (x_1, y_1), (x_2, y_2) we have

$$(x_1 - x_2)\cos t + (y_1 - y_2)\sin t = f(t) + f''(t)$$

and $$(x_1 - x_2)\sin t - (y_1 - y_2)\cos t = 0,$$

whence, adding the squares of the two equations, the radius of curvature ρ is given by

$$\rho^2 = \{f(t) + f''(t)\}^2.$$

Since $L(t) = 0$ is a tangent to its envelope, it follows that if ψ is the inclination of the tangent, then $\tan\psi = -\cot t = \tan(t - \tfrac{1}{2}\pi)$, and therefore ψ differs from t by $(n+\tfrac{1}{2})\pi$ for some n, whence, on the envelope,

$$\rho = \frac{ds}{d\psi} = \frac{ds}{dt},$$

and therefore $$\frac{ds}{dt} = |f(t) + f''(t)|.$$

If $f(t)+f''(t)$ is non-negative, then

$$s = \int \{f(t)+f''(t)\}\,dt = f'(t)+\int f(t)\,dt,$$

and if $f(t)+f''(t)$ is negative then

$$s = -\int \{f(t)+f''(t)\}\,dt = -f'(t)-\int f(t)\,dt,$$

so that, provided $f(t)+f''(t)$ is of constant sign, $s = |f'(t)+\int f(t)\,dt|$, since s is necessarily positive.

5. If $a = a(t)$, $b = b(t)$, and $R = R(t)$, and if s is the length of the arc of the curve $x = a(t)$, $y = b(t)$, show that the family of circles

$$(x-a)^2+(y-b)^2 = R^2$$

admits an envelope if and only if $(dR/ds)^2 \leqslant 1$.

The envelope, if any, satisfies the equations

$$C(t) \equiv (x-a)^2+(y-b)^2-R^2 = 0,$$

$$\tfrac{1}{2}C'(t) \equiv (x-a)a'+(y-b)b'-RR' = 0.$$

The distance of the line $C'(t) = 0$ from the centre of the circle $C(t) = 0$ is

$$\frac{RR'}{|\sqrt{(a'^2+b'^2)}} \qquad R\frac{dR}{ds} \qquad \text{since} \quad s' = \frac{ds}{dt} = \sqrt{(a'^2+b'^2)}.$$

Thus $C'(t) = 0$ intersects $C(t) = 0$ if and only if $|dR/ds| \leqslant 1$.

If $|dR/ds| < 1$, the line $C'(t) = 0$ meets the circle $C(t) = 0$ at two distinct points, and so the envelope of the family of circles touches each circle at two distinct points.

If $|dR/ds| = 1$, the line $C'(t) = 0$ touches the circle $C(t) = 0$, and therefore also touches the envelope; the tangent to the locus of the centre of $C(t) = 0$ is parallel to $x/a' = y/b'$ and therefore perpendicular to $C'(t) = 0$, so that the tangent to the locus of the centre of the circle $C(t) = 0$ is normal to the envelope of the circle. Accordingly the locus of the centre of $C(t) = 0$ is the evolute of the envelope of $C(t) = 0$, and therefore $C(t) = 0$ is the circle of curvature of the envelope of the family of circles.

6. Show that the family of circles $(x-\cos t)^2+(y-\sin t)^2 = t^2$ admits an envelope which is *not* a locus of the limit of points of intersection of two circles of the family.

The locus of the centres of the circles of the family is $x = \cos t$, $y = \sin t$ so that the arc length of this locus, measured from $t = 0$, is t. The radius of the circle $C(t) \equiv (x-\cos t)^2+(y-\sin t)^2-t^2 = 0$

is $R = t$, so that $dR/ds = dt/dt = 1$. Hence, by the previous example, the family of circles admits an envelope, which has contact of the second order with each circle. The distance between the centres of the circles $C(t) = 0$ and $C(T) = 0$, $T > t > 0$, is

$$\{(\cos T-\cos t)^2+(\sin T-\sin t)^2\}^{\frac{1}{2}} = 2\sin\frac{T-t}{2} < T-t,$$

the difference between the radii, and so the circle $C(t) = 0$ is completely contained inside the circle $C(T) = 0$.

This example illustrates a particular case of Theorem 16.61.

16.6. In the neighbourhood of a non-singular point on a family of curves $f(x, y, a) = 0$ at least one of the derivatives $\partial f/\partial x$, $\partial f/\partial y$ is different from zero. Suppose the latter, then by Theorem 15.85, in this neighbourhood, the equation $f(x, y, a) = 0$ has a differentiable solution $y = \phi(x, a)$. We shall show that, in this neighbourhood, the necessary and sufficient condition for the envelope of the family $f(x, y, a) = 0$ to have contact of the nth order exactly with each curve of the family is that

$$\frac{\partial\phi}{\partial a} = \frac{\partial^2\phi}{\partial a^2} = \dots = \frac{\partial^n\phi}{\partial a^n} = 0 \quad \text{and} \quad \frac{\partial^{n+1}\phi}{\partial a^{n+1}} \neq 0.$$

For if these conditions are satisfied in some neighbourhood then from $\partial^n\phi/\partial a^n = 0$ and $\partial^{n+1}\phi/\partial a^{n+1} \neq 0$ it follows from Theorem 15.85 that we can determine a as a unique function of x, say $a = a(x)$, satisfying $\partial^n\phi/\partial a^n = 0$, and since this solution is unique, and since $\partial^r\phi/\partial a^r = 0$, $r = 1, 2,\dots, n$, simultaneously, therefore $a = a(x)$ satisfies all the equations $\partial^r\phi/\partial a^r = 0$; in particular $a = a(x)$ satisfies $\partial\phi/\partial a = 0$ and so the family $y = \phi(x, a)$ admits an envelope given by the equation $y = \phi\{x, a(x)\}$. Provided this envelope does not coincide with a curve of the family, $a(x)$ is not constant in the neighbourhood under consideration, and hence there is a region in which $a'(x) \neq 0$. But, differentiating the equation $\dfrac{\partial^n}{\partial a^n}\phi\{x, a(x)\} = 0$ with respect to x, we find

$$\frac{\partial^{n+1}\phi}{\partial a^n\partial x}+\frac{\partial^{n+1}\phi}{\partial a^{n+1}}a'(x) = 0,$$

and since $\partial^{n+1}\phi/\partial a^{n+1}$ and $a'(x)$ are both different from zero, therefore $\dfrac{\partial}{\partial x}\left(\dfrac{\partial^n\phi}{\partial a^n}\right)$ is not zero, and so, again by Theorem 15.85, the

equation $\partial^n\phi/\partial a^n = 0$ is solvable for x in terms of a, giving $x = x(a)$. Accordingly the equation of the envelope is $y = \phi\{x(a), a\}, x = x(a)$. Transforming the independent variable x to the new variable a, by the transformation $x = x(a)$, the equations of the envelope, and the curve of the family which has the parameter a_0, take the forms

$$y = \phi\{x(a), a\} \quad \text{and} \quad y = \phi\{x(a), a_0\}.$$

These curves meet at the point $a = a_0$, $y = \phi\{x(a_0), a_0\}$ and have contact of the nth order exactly in virtue of the conditions $\partial^r\phi/\partial a^r = 0$, $r = 1, 2,..., n$, and $\partial^{n+1}\phi/\partial a^{n+1} \neq 0$, at the point $a = a_0$, since by Taylor's theorem we can find α in (a_0, a), such that

$$\begin{aligned}
&\phi\{x(a), a_0\}-\phi\{x(a), a\} \\
&\quad = \phi\{x(a), a+(a_0-a)\}-\phi\{x(a), a\} \\
&\quad = (a_0-a)\frac{\partial}{\partial a}\phi\{x(a), a\}+\frac{(a_0-a)^2}{2!}\frac{\partial^2}{\partial a^2}\phi\{x(a), a\}+...+ \\
&\qquad +\frac{(a_0-a)^n}{n!}\frac{\partial^n}{\partial a^n}\phi\{x(a), a\}+\frac{(a_0-a)^{n+1}}{(n+1)!}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), \alpha\} \\
&\quad = \frac{(a_0-a)^{n+1}}{(n+1)!}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), \alpha\}
\end{aligned}$$

whence $$\lim_{a\to a_0}\frac{\phi\{x(a), a_0\}-\phi\{x(a), a\}}{(a_0-a)^n} = 0.$$

Conversely, if the family $y = \phi(x, a)$ admits an envelope, then the equations $y = \phi(x, a)$, $\partial\phi/\partial a = 0$ have a solution $x = x(a)$, $y = \phi\{x(a), a\}$. Transforming the variables from x, y to a, y the equation of the envelope becomes $y = \phi\{x(a), a\}$, and the curve with parameter a_0, $y = f(x, a_0)$, becomes $y = \phi\{x(a), a_0\}$. Since the envelope has contact of the nth order exactly, with each member of the family, therefore

$$\lim_{a\to a_0}\frac{\phi\{x(a), a\}-\phi\{x(a), a_0\}}{(a-a_0)^r} = 0, \quad 0 \leqslant r \leqslant n, \tag{i}$$

for all a_0 in the region under consideration.

But $\phi\{x(a), a\}-\phi\{x(a), a_0\} = (a-a_0)\dfrac{\partial}{\partial a}\phi\{x(a), \alpha\}$, for a certain α in (a_0, a), whence, taking $r = 1$, we find that $\dfrac{\partial}{\partial a}\phi\{x(a), \alpha\} \to 0$ as

$a \to a_0$, i.e. $\dfrac{\partial}{\partial a}\phi\{x(a_0), a_0\} = 0$, since $\alpha \to 0$ as $a \to a_0$. Since this holds for all a_0 in some region we may write $\dfrac{\partial}{\partial a}\phi\{x(a), a\} = 0$. Furthermore, if we have proved, step by step, that

$$\frac{\partial^r}{\partial a^r}[\phi\{x(a), a\}] = 0, \quad \text{for} \quad 0 \leqslant r \leqslant k < n,$$

then

$$\phi\{x(a), a_0\} - \phi\{x(a), a\} = \phi\{x(a), a+(a_0-a)\} - \phi\{x(a), a\}$$
$$= \frac{(a_0-a)^{k+1}}{(k+1)!}\frac{\partial^{k+1}}{\partial a^{k+1}}\phi\{x(a), \alpha\} \quad \text{for some } \alpha \text{ in } (a_0, a)$$

whence, taking $r = k+1$ in (i), it follows that

$$\frac{\partial^{k+1}}{\partial a^{k+1}}\phi\{x(a), \alpha\} \to 0 \quad \text{as} \quad a \to a_0,$$

i.e. $\dfrac{\partial^{k+1}}{\partial a^{k+1}}\phi\{x(a_0), a_0\} = 0$. Since this holds at all points a_0 in our region, we have $\dfrac{\partial^{k+1}}{\partial a^{k+1}}\phi\{x(a), a\} = 0$, and therefore

$$\frac{\partial^r}{\partial a^r}\phi\{x(a), a\} = 0, \quad 0 \leqslant r \leqslant n.$$

16.61. If the curve C, with equations $x = x(a)$, $y = y(a)$ is the envelope of the family of curves $y = \phi(x, a)$, and if C has contact of the nth order exactly with each curve of the family, then for all a_0 the point $x = x(a_0)$, $y = y(a_0)$, where C touches the curve of parameter a_0, is the limit of the point of intersection of the curves of parameters a_0, a_1 as $a_1 \to a_0$, *if and only if* n is an odd number.

Proof. Since

$$\frac{\partial^r}{\partial a^r}\phi\{x(a), a\} = 0, \quad 1 \leqslant r \leqslant n,$$

therefore we can find θ in [0, 1] such that

$$\phi\{x(a), a+t\} = \phi\{x(a), a\} + t\frac{\partial \phi}{\partial a} + \frac{t^2}{2!}\frac{\partial^2 \phi}{\partial a^2} + \dots +$$
$$+ \frac{t^n}{n!}\frac{\partial^n}{\partial a^n}\phi\{x(a), a\} + \frac{t^{n+1}}{(n+1)!}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), a+\theta t\}$$
$$= \frac{t^{n+1}}{(n+1)!}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), a+\theta t\},$$

whence

$$\begin{aligned}\phi\{x(a), a_1\}-\phi\{x(a), a_0\} &= \phi\{x(a), a+(a_1-a)\}-\phi\{x(a), a+(a_0-a)\} \\ &= \frac{1}{(n+1)!}\left\{(a_1-a)^{n+1}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), \alpha\}-(a_0-a)^{n+1}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), \beta\}\right\},\end{aligned}$$

where α lies in (a, a_1) and β in (a, a_0).

Suppose first that n is odd. Then $\phi\{x(a), a_1\}-\phi\{x(a), a_0\}$ takes the value $(a_1-a_0)^{n+1}\dfrac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a_0), \alpha_0\}$ when $a = a_0$, and the value $-(a_1-a_0)^{n+1}\dfrac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a_1), \beta_1\}$ when $a = a_1$, where α_0 is the value of α at $a = a_0$, and β_1 is the value of β at $a = a_1$. Since $\dfrac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), a\}$ is of constant sign in a neighbourhood of $a = a_0$ it follows that $\phi\{x(a), a_1\}-\phi\{x(a), a_0\}$ changes sign as a varies from a_0 to a_1, and therefore this difference vanishes at a point $a = a_0'$ between a_0 and a_1. Thus the two curves $y = \phi\{x(a), a_0\}$, $y = \phi\{x(a), a_1\}$ intersect at the point $x = x(a_0')$, and since a_0' lies between a_0 and a_1, therefore $x(a_0') \to x(a_0)$ when $a_1 \to a_0$.

The proof that the curves do not intersect when n is even is rather more difficult.

We consider first values of a *between* a_0 and a_1.

For such values of a, $(a_1-a)^{n+1}$ and $(a_0-a)^{n+1}$ have opposite signs and therefore

$$(a_1-a)^{n+1}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), \alpha\} \quad \text{and} \quad -(a_0-a)^{n+1}\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), \beta\}$$

have the same sign and do not vanish, so that

$$\phi\{x(a), a_1\}-\phi\{x(a), a_0\}$$

is of constant sign for an a between a_0 and a_1 and does not vanish.

When a lies *outside* the interval (a_0, a_1), we use Theorem 12.52.

Let $P(t) = \phi\{x(a), a+t\}$, $Q(t) = t^{n+1}$, and $u = a_1-a$, $v = a_0-a$. Since a lies outside (a_0, a_1) therefore u and v have the same sign, and it is readily verified that the remaining conditions of 12.52 are satisfied. Hence

$$\begin{aligned}\phi\{x(a), a_1\}-\phi\{x(a), a_0\} &= P(u)-P(v) \\ &= \frac{(a_1-a)^{n+1}-(a_0-a)^{n+1}}{(n+1)!}\,\frac{\partial^{n+1}}{\partial a^{n+1}}\phi\{x(a), a+c_{n+1}\}.\end{aligned}$$

The differences a_1-a, a_0-a have the same sign; since

$$(a_1-a)^{n+1}-(a_0-a)^{n+1} = (a_1-a_0)\sum (a_1-a)^r(a_0-a)^{n-r},$$

and $$\sum (a_1-a)^r(a_0-a)^{n-r} = (a_0-a)^n \sum \left(\frac{a_1-a}{a_0-a}\right)^r > 0,$$

therefore $(a_1-a)^{n+1}-(a_0-a)^{n+1}$ cannot vanish for an a outside (a_0, a_1).

Accordingly $\phi\{x(a), a_1\}-\phi\{x(a), a_0\}$ does not vanish for any value of a, and so, when n is even, the curves $y = \phi\{x(a), a_1\}$, $y = \phi\{x(a), a_0\}$ do not intersect.

XVII

DOUBLE INTEGRALS

REDUCTION OF DOUBLE INTEGRALS TO REPEATED INTEGRALS. GREEN'S THEOREM. TRANSFORMATION OF DOUBLE INTEGRALS

17. A function $f(x, y)$ is continuous in a (two-dimensional) interval $(a, A)(b, B)$. Since $f(x, y)$ is x-continuous in (a, A) then $\int_a^A f(x, y)\,dx$ exists, and since $f(x, y)$ is continuous, the integral is y-continuous, and so $\int_b^B \left\{\int_a^A f(x, y)\,dx\right\} dy$ exists. Similarly $\int_a^A \left\{\int_b^B f(x, y)\,dy\right\} dx$ exists.

17.1. We prove that

$$\int_b^B \left\{\int_a^A f(x, y)\,dx\right\} dy = \int_a^A \left\{\int_b^B f(x, y)\,dy\right\} dx.$$

Corresponding to any k we can divide $(a, A)(b, B)$ into sub-intervals $(a_r, a_{r+1})(b_s, b_{s+1})$, $r = 0, 1, 2,..., m$, $s = 0, 1, 2,..., n$, where m and n depend upon k, and $a_0 = a$, $a_{m+1} = A$, $b_0 = b$, $b_{n+1} = B$, such that

$$f(x, y)-f(X, Y) = 0(k)$$

for any two points (x, y), (X, Y) in the same sub-interval. Then

$$\int_a^A f(x, y)\,dx = \sum_0^m f(a_r, y)(a_{r+1}-a_r)+(A-a)0(k)$$

and so

$$\int_b^B \left\{\int_a^A f(x, y)\,dx\right\} dy$$

$$= \sum_0^m \left[(a_{r+1}-a_r)\left\{\sum_0^n (b_{s+1}-b_s)f(a_r, b_s)+(B-b)0(k)\right\}\right]+$$
$$+(B-b)(A-a)0(k),$$
$$= \sum_{r,s} \rho_{r,s} f(a_r, b_s)+2\rho\,.\,0(k),$$

where $\rho_{r,s}$ denotes the area of the interval $(a_r, a_{r+1})(b_s, b_{s+1})$, ρ the area of $(a, A)(b, B)$ and the summation is taken over all pairs r, s, with $r = 0, 1,..., m$ and $s = 0, 1, 2,..., n$.

Similarly

$$\int_a^A \left\{ \int_b^B f(x, y)\, dy \right\} dx = \sum_{r,s} \rho_{r,s} f(a_r, b_s) + 2\rho . O(k),$$

and therefore

$$\int_b^B \left\{ \int_a^A f(x, y)\, dx \right\} dy - \int_a^A \left\{ \int_b^B f(x, y)\, dy \right\} dx = 4\rho . O(k),$$

and since this is true for any k,

$$\int_b^B \left\{ \int_a^A f(x, y)\, dx \right\} dy = \int_a^A \left\{ \int_b^B f(x, y)\, dy \right\} dx.$$

The integrals

$$\int_b^B \left\{ \int_a^A f(x, y)\, dx \right\} dy \quad \text{and} \quad \int_a^A \left\{ \int_b^B f(x, y)\, dy \right\} dx$$

are called *repeated* integrals. We have seen that the order of repeated integration of a continuous function is immaterial, so that the integral depends only upon the interval $(a, A)(b, B)$. If this interval is called R, it is customary to denote the common value of the repeated integrals by

$$\int_R f(x, y)\, dR.$$

The repeated integrals themselves are abbreviated to

$$\int_b^B \int_a^A f(x, y)\, dxdy \quad \text{and} \quad \int_a^A \int_b^B f(x, y)\, dydx \quad \text{respectively.}$$

17.11. Given any p, we can divide the interval R into sub-intervals $i_{r,s}$ such that, if (α_r, β_s) is *any* point in $i_{r,s}$, and $\rho_{r,s}$ is the area of $i_{r,s}$, then

$$\int_R f(x, y)\, dR = \sum_{r,s} \rho_{r,s} f(\alpha_r, \beta_s) + O(p).$$

For, by 17.1,

$$\int_R f(x, y)\, dR = \sum_{r,s} \rho_{r,s} f(a_r, b_s) + 2\rho . O(k),$$

and $f(\alpha_r, \beta_s) - f(a_r, b_s) = O(k)$, so that

$$\begin{aligned}\int_R f(x, y)\, dR - \sum_{r,s} \rho_{r,s} f(\alpha_r, \beta_s) &= \rho . O(k) + 2\rho . O(k)\\ &= 3\rho . O(k) = O(p)\end{aligned}$$

by a suitable choice of k.

17.2. Integrals along an elementary closed curve

Given the numbers α, β, γ, δ, α_1, α_2, β_1, β_2, γ_1, γ_2, δ_1 and δ_2, satisfying

$$\alpha < \beta, \quad \gamma < \delta, \quad \gamma \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \delta, \quad \gamma \leqslant \beta_1 \leqslant \beta_2 \leqslant \delta,$$
$$\alpha \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \beta, \quad \alpha \leqslant \delta_1 \leqslant \delta_2 \leqslant \beta,$$

we define an *elementary* (or *simple*) closed curve as follows:

C consists of the line segment $x = \alpha$ from $y = \alpha_1$ to $y = \alpha_2$; an arc of a curve whose equation is $y = \lambda(x)$ (or $x = \lambda^{-1}(y)$) from $x = \alpha$ to $x = \delta_1$, where $\lambda(x)$ is a continuous steadily increasing function; the segment $y = \delta$, $\delta_1 \leqslant x \leqslant \delta_2$; the arc $y = \mu(x)$, $\delta_2 \leqslant x \leqslant \beta$, where $\mu(x)$ is continuous and steadily decreasing; the segment $x = \beta$, $\beta_1 \leqslant y \leqslant \beta_2$; the arc $y = \nu(x)$, $\gamma_2 \leqslant x \leqslant \beta$, where $\nu(x)$ is continuous and steadily increasing; the segment $y = \gamma$, $\gamma_1 \leqslant x \leqslant \gamma_2$; the arc $y = \rho(x)$, $\alpha \leqslant x \leqslant \gamma_1$, where $\rho(x)$ is continuous and steadily decreasing. (Note that $\lambda(\alpha) = \alpha_2$, $\lambda(\delta_1) = \delta = \mu(\delta_2)$, etc.)

It is readily verified that a circle, an ellipse, a rectangle, or a triangle is an elementary closed curve according to this definition.

The rectangle $(\alpha, \beta)(\gamma, \delta)$ is called the *bounding rectangle* of the curve C.

If we define the functions $x_1(x)$, $x_2(x)$ by the conditions

$$\begin{aligned} x_1(x) &= \rho(x), & \alpha &\leqslant x \leqslant \gamma_1 \qquad & x_2(x) &= \lambda(x), & \alpha &\leqslant x \leqslant \delta_1 \\ &= \gamma, & \gamma_1 &\leqslant x \leqslant \gamma_2 & &= \delta, & \delta_1 &\leqslant x \leqslant \delta_2 \\ &= \nu(x), & \gamma_2 &\leqslant x \leqslant \beta; & &= \mu(x), & \delta_2 &\leqslant x \leqslant \beta, \end{aligned}$$

then $x_1(x)$, $x_2(x)$ are continuous functions and any line $x = x^*$, $\alpha < x^* < \beta$, meets C in exactly two points $y = x_1(x^*)$, $y = x_2(x^*)$, where $x_1(x^*) < x_2(x^*)$.

Similarly if we define

$$\begin{aligned} y_1(y) &= \rho^{-1}(y), & \gamma &\leqslant y \leqslant \alpha_1 \qquad & y_2(y) &= \nu^{-1}(y), & \gamma &\leqslant y \leqslant \beta \\ &= \alpha, & \alpha &\leqslant y \leqslant \alpha_2 & &= \beta, & \beta_1 &\leqslant y \leqslant \beta_2 \\ &= \lambda^{-1}(y), & \alpha_2 &\leqslant y \leqslant \delta; & &= \mu^{-1}(y), & \beta_2 &\leqslant y \leqslant \delta, \end{aligned}$$

then $y_1(y)$, $y_2(y)$ are continuous and any line $y = y^*$, $\gamma < y^* < \delta$, meets C in exactly two points $y_1(y^*)$, $y_2(y^*)$, where $y_1(y^*) < y_2(y^*)$.

If $\alpha < x < \beta$ and $x_1(x) < y < x_2(x)$ *or* $\gamma < y < \delta$ and

$$y_1(y) < x < y_2(y)$$

then the point (x, y) is said to lie *inside* the curve C. If (x, y) is not inside or on C, then it is said to lie *outside* C.

It follows that if we remove from the bounding rectangle $(\alpha, \beta)(\gamma, \delta)$ the four rectangles $(\alpha, \delta_1)(\alpha_2, \delta)$, $(\delta_2, \beta)(\beta_2, \delta)$, $(\gamma_2, \beta)(\gamma, \beta_1)$, and $(\alpha, \gamma_1)(\gamma, \alpha_1)$ there remain only points which are inside or on the curve C.

Furthermore, each of the rectangles $(\alpha, \delta_1)(\alpha_2, \delta)$, etc., is divided into two parts by an arc of the curve C. For the equation of the arc in $(\alpha, \delta_1)(\alpha_2, \delta)$ is $y = x_2(x)$ and all points (x, y) such that $\alpha_2 \leqslant y < x_2(x)$ are *inside* the curve, and all points (x, y) such that $x_2(x) < y \leqslant \delta$ are *outside* the curve.

If $P(x, y)$ is continuous on C, we define the *anti-clockwise integrals along C*,

$$\int_C P(x, y)\, dx = \int_\alpha^\beta \{P(x, x_1) - P(x, x_2)\}\, dx$$

and

$$\int_C P(x, y)\, dy = \int_\gamma^\delta \{P(y_2, y) - P(y_1, y)\}\, dy,$$

and the *clockwise integrals*

$$\int_C^* P(x, y)\, dx = \int_\alpha^\beta \{P(x, x_2) - P(x, x_1)\}\, dx,$$

$$\int_C^* P(x, y)\, dy = \int_\gamma^\delta \{P(y, y_1) - P(y, y_2)\}\, dy.$$

17.21. *If a bounded function $f(x, y)$ is continuous inside and on a closed curve C which is bounded by the rectangle $(\alpha, \beta)(\gamma, \delta)$, and is continuous inside the rectangle and outside C, then*

$$\int_\alpha^\beta f(x, y)\, dx$$

exists and is y-continuous in (γ, δ).

Observe first that $f(x, y)$ is not necessarily continuous throughout the rectangle $(\alpha, \beta)(\gamma, \delta)$ since, for instance, a function which takes the value unity inside and on C, but is zero outside C, is continuous inside and on C, and continuous outside C, but is not continuous throughout the rectangle.

Let $\{x^*, x_1(x^*)\}$, $\{x^*, x_2(x^*)\}$ be the points where $x = x^*$, $\alpha \leqslant x^* \leqslant \beta$, meets the curve, and $\{y_1(y^*), y^*\}$, $\{y_2(y^*), y^*\}$ be the points where $y = y^*$, $\gamma \leqslant y^* \leqslant \delta$, meets the curve.

Since $y_1(y)$ and $y_2(y)$ are continuous we can choose n_k so that $y_1(Y)-y_1(y) = 0(k)$, $y_2(Y)-y_2(y) = 0(k)$, provided $Y-y = 0(n_k)$ and both y and Y lie between γ and δ.

Furthermore, since $f(x, y)$ is continuous inside C, and continuous outside C, we may suppose n_k chosen so that $|f(x, Y)-f(x, y)| < 1/k$, provided $Y-y = 0(n_k)$ and (x, y), (x, Y) both lie inside C, or both lie outside. Now $f(x, y)$ is x-continuous in each of the intervals $(\alpha, y_1]$, (y_1, y_2), and $[y_2, \beta)$, for these intervals lie wholly inside or wholly outside C, $\gamma \leqslant y \leqslant \delta$, and $f(x, y)$ is bounded in (α, β), and therefore each of the integrals $\int_\alpha^{y_1} f(x, y)\,dx$, $\int_{y_1}^{y_2} f(x, y)\,dx$, and $\int_{y_2}^{\beta} f(x, y)\,dx$ exists.

Hence if we *define*

$$\int_\alpha^\beta f(x, y)\,dx = \int_\alpha^{y_1} f(x, y)\,dx + \int_{y_1}^{y_2} f(x, y)\,dx + \int_{y_2}^{\beta} f(x, y)\,dx,$$

then $\int_\alpha^\beta f(x, y)\,dx$ exists.

Consider

$$\int_\alpha^\beta f(x, y)\,dx - \int_\alpha^\beta f(x, Y)\,dx = \int_\alpha^\beta \{f(x, y)-f(x, Y)\}\,dx$$

$$= \int_\alpha^{y_1-1/k} + \int_{y_2+1/k}^{\beta} + \int_{y_1+1/k}^{y_2-1/k} + \int_{y_1-1/k}^{y_1+1/k} + \int_{y_2-1/k}^{y_2+1/k} \{f(x, y)-f(x, Y)\}\,dx.$$

If $Y-y = 0(n_k)$, both $y_1(y)$ and $y_1(Y)$ lie between

$$y_1(y)-1/k \quad \text{and} \quad y_1(y)+1/k,$$

for $|y_1(Y)-y_1(y)| < 1/10^k < 1/k$, and similarly for $y_2(y)$ and $y_2(Y)$, and therefore when x lies in $(\alpha, y_1-1/k)$ both (x, y) and (x, Y) are outside C so that $|f(x, y)-f(x, Y)| < 1/k$. The same reasoning shows that the inequality $|f(x, y)-f(x, Y)| < 1/k$ holds also when x lies in either of the intervals $(y_2+1/k, \beta)$, $(y_1+1/k, y_2-1/k)$.

In

$$(y_1-1/k,\ y_1+1/k) \quad \text{and} \quad (y_2-1/k,\ y_2+1/k),$$

$$|f(x, y)-f(x, Y)| < 2M,$$

where M is a bound of $f(x, y)$ in $(\alpha, \beta)(\gamma, \delta)$.

Hence

$$\left|\int_{\alpha}^{\beta} f(x,y)\,dx - \int_{\alpha}^{\beta} f(x,Y)\,dx\right| < \{(y_1-\alpha)+(\beta-y_2)+(y_2-y_1)\}/k+8M/k$$
$$= \{8M+\beta-\alpha\}/k,$$

which proves that $\int_{\alpha}^{\beta} f(x, y)\,dx$ is y-continuous in (γ, δ).

17.22. If $f(x, y)$ is continuous inside and on a closed curve C which is bounded by the rectangle $(\alpha, \beta)(\gamma, \delta)$ and is continuous outside C, then

$$\int_{\gamma}^{\delta} f(x, y)\,dy$$

exists and is continuous for x in (α, β).

Proof the same as in 17.21.

17.23. It follows from 17.21 and 17.22 that both the repeated integrals

$$\int_{\gamma}^{\delta}\int_{\alpha}^{\beta} f(x, y)\,dxdy, \qquad \int_{\alpha}^{\beta}\int_{\gamma}^{\delta} f(x, y)\,dydx$$

exist. We prove next that these repeated integrals are equal.

We observe first that if we divide a rectangle $(\alpha, \beta)(\gamma, \delta)$ into four rectangles by the lines $x = \lambda$ and $y = \mu$ then a repeated integral over $(\alpha, \beta)(\gamma, \delta)$ is the sum of the repeated integrals over the rectangles $(\alpha, \lambda)(\gamma, \mu)$, $(\alpha, \lambda)(\mu, \delta)$, $(\lambda, \beta)(\gamma, \mu)$, and $(\lambda, \beta)(\mu, \delta)$, for

$$\int_{\gamma}^{\beta} f\,dx = \int^{\lambda} + \int^{\beta} f\,dx$$

and

$$\int^{\delta} \phi\,dy = \int^{\mu} + \int^{\delta} \phi\,dy$$

and so

$$\int_{\gamma}^{\delta}\int_{\alpha}^{\beta} f\,dx = \int_{\gamma}^{\delta}\left\{\int_{\alpha}^{\lambda} + \int_{\lambda}^{\beta} f\,dx\right\}dy = \int_{\gamma}^{\delta}\int_{\alpha}^{\lambda} + \int_{\gamma}^{\delta}\int_{\lambda}^{\beta} f\,dxdy$$
$$= \int_{\gamma}^{\mu}\int_{\alpha}^{\lambda} + \int_{\gamma}^{\mu}\int_{\lambda}^{\beta} + \int_{\mu}^{\delta}\int_{\alpha}^{\lambda} + \int_{\mu}^{\delta}\int_{\lambda}^{\beta} f\,dxdy.$$

Consequently, if a rectangle R is divided into any number of rectangles $R_1, R_2,\ldots, R_k$ then a repeated integral over R is the

sum of the repeated integrals over each of the rectangles $R_1, R_2,\ldots, R_k$. We have seen in 17.2 that if we remove from the bounding rectangle $(\alpha, \beta)(\gamma, \delta)$ the four rectangles $(\alpha, \delta_1)(\alpha_2, \delta)$, $(\delta_2, \beta)(\beta_2, \delta)$, $(\gamma_2, \beta)(\gamma, \beta_1)$, and $(\alpha, \gamma_1)(\gamma, \alpha_1)$, the remainder of the bounding rectangle is entirely inside or on C; this remainder may be divided into rectangles (in several ways, in fact) in each of which $f(x, y)$ is continuous, so that the repeated integrals over these rectangles are equal (by 17.1). It remains to show that the repeated integrals over such a rectangle as $(\alpha, \delta_1)(\alpha_2, \delta)$ are also equal.

The arc of C contained in $(\alpha, \delta_1)(\alpha_2, \delta)$ is $y = x_2(x)$, where $x_2(x)$ is a steadily increasing continuous function.

Divide (α, δ_1) into n equal parts by the points

$$\alpha = p_0, p_1, p_2,\ldots, p_n = \delta_1,$$

so that $p_{r+1}-p_r = (\delta_1-\alpha)/n$ and let $x_2(p_k) = q_k$ for $k = 0, 1, 2,\ldots, n$.

The lines $x = p_r$, $y = q_s$, r, $s = 0, 1, 2,\ldots, n$, divide the rectangle $(\alpha, \delta_1)(\alpha_2, \delta)$ into n^2 rectangles; of these the n rectangles $(p_0, p_1)(q_0, q_1)$, $(p_1, p_2)(q_1, q_2)$, $(p_2, p_3)(q_2, q_3),\ldots, (p_{n-1}, p_n)(q_{n-1}, q_n)$ contain all the points of the arc (for if $p_r \leqslant x \leqslant p_{r+1}$ then $q_r \leqslant x_2(x) \leqslant q_{r+1}$, since $x_2(x)$ is steadily increasing), and so each of the remaining rectangles is either completely inside C, or completely outside C. Thus the repeated integrals over the rectangle $(p_r, p_{r+1})(q_s, q_{s+1})$ are equal provided r and s are unequal, and there remains to consider only the sum of the integrals over the rectangles $(p_r, p_{r+1})(q_r, q_{r+1})$, $r = 0, 1, 2,\ldots, n$.

Since $f(x, y)$ is bounded by M

$$\left|\int_{q_r}^{q_{r+1}}\int_{p_r}^{p_{r+1}} f(x, y)\,dxdy\right| < \int_{q_r}^{q_{r+1}} M(p_{r+1}-p_r)\,dy = M(p_{r+1}-p_r)(q_{r+1}-q_r)$$

and so

$$\sum_{0}^{n-1}\int_{q_r}^{q_{r+1}}\int_{p_r}^{p_{r+1}} f(x, y)\,dxdy < M\sum_{0}^{n-1}(p_{r+1}-p_r)(q_{r+1}-q_r)$$

$$= M\{(\delta_1-\alpha)/n\}\sum_{0}^{n-1}(q_{r+1}-q_r)$$

$$= M(\delta_1-\alpha)(\delta-\alpha_2)/n = MA/n,$$

where A is the area of $(\alpha, \delta_1)(\alpha_2, \delta)$, and similarly

$$\left|\sum_{0}^{n-1}\int_{p_r}^{p_{r+1}}\int_{q_r}^{q_{r+1}} f(x, y)\,dydx\right| < MA/n.$$

Hence

$$\int_{\alpha_2}^{\delta}\int_{\alpha}^{\delta_1} f(x,y)\,dxdy - \int_{\alpha}^{\delta_1}\int_{\alpha_2}^{\delta} f(x,y)\,dydx$$

$$= \left|\sum_{0}^{n-1}\int_{q_r}^{q_{r+1}}\int_{p_r}^{p_{r+1}} f(x,y)\,dxdy - \sum_{0}^{n-1}\int_{p_r}^{p_{r+1}}\int_{q_r}^{q_{r+1}} f(x,y)\,dydx\right|$$

$$< 2MA/n.$$

But we may choose n as great as we please, and therefore

$$\int_{\alpha_2}^{\delta}\int_{\alpha}^{\delta_1} f\,dxdy = \int_{\alpha}^{\delta_1}\int_{\alpha_2}^{\delta} f\,dydx,$$

which completes the proof.

17.231. By means of Theorem 17.23 we can give a simpler proof of Theorem 15.91, and establish the result of that theorem under rather less stringent conditions.

Let $f(x, y)$ be continuous and y-differentiable in $(a, b)(c, d)$, and write

$$\phi(y) = \int_{a}^{b} f(x,y)\,dx, \qquad \psi(y) = \int_{a}^{b} f_y(x,y)\,dx.$$

Then, for a t in (c, d),

$$\int_{c}^{t} \psi(y)\,dy = \int_{c}^{t}\left(\int_{a}^{b} f_y(x,y)\,dx\right)dy = \int_{a}^{b}\left(\int_{c}^{t} f_y(x,y)\,dy\right)dx$$

$$= \int_{a}^{b} \{f(x,t)-f(x,c)\}\,dx = \phi(t)-\phi(c),$$

whence, differentiating with respect to t,

$$\psi(t) = \phi'(t), \quad \text{for any } t \text{ in } (c, d),$$

i.e.

$$\int_{a}^{b} f_y(x,y)\,dx = \frac{d}{dy}\int_{a}^{b} f(x,y)\,dx.$$

17.3. The integral of a continuous function $f(x, y)$ over the interior of a closed curve C

Let $(\alpha,\beta)(\gamma,\delta)$ be the rectangle bounding C, and let $f_1(x,y) = f(x,y)$ at any point (x, y) inside or on C, and $f_1(x, y) = 0$ outside C; then $f_1(x, y)$ is continuous inside C and continuous outside C. Hence

both the integrals $\int\limits_{\alpha}^{\beta}\int\limits_{\gamma}^{\delta} f_1(x, y)\, dydx$ and $\int\limits_{\gamma}^{\delta}\int\limits_{\alpha}^{\beta} f_1(x, y)\, dxdy$ exist and are equal. To any x between α and β correspond points of C, $\{x, x_1(x)\}$, $\{x, x_2(x)\}$, and to y between γ and δ correspond points of C, $\{y_1(y), y\}$, $\{y_2(y), y\}$. Since $f_1(x, y) = f(x, y)$ for $y_1(y) \leqslant x \leqslant y_2(y)$ and $f_1(x, y) = 0$ for $x < y_1(y)$ or $x > y_2(y)$, therefore

$$\int\limits_{\alpha}^{\beta} f_1(x, y)\, dx = \int\limits_{y_1}^{y_2} f(x, y)\, dx,$$

and since $f_1(x, y) = f(x, y)$ for $x_1(x) \leqslant y \leqslant x_2(x)$ and $f_1(x, y) = 0$ for $y < x_1(x)$ or $y > x_2(x)$, therefore

$$\int\limits_{\gamma}^{\delta} f_1(x, y)\, dy = \int\limits_{x_1}^{x_2} f(x, y)\, dy.$$

Hence both the integrals

$$\int\limits_{\gamma}^{\delta}\int\limits_{y_1}^{y_2} f(x, y)\, dxdy, \qquad \int\limits_{\alpha}^{\beta}\int\limits_{x_1}^{x_2} f(x, y)\, dydx$$

exist and are equal; their common value is denoted by

$$\int\limits_{C} f(x, y)\, dxdy,$$

which is called the integral of $f(x, y)$ over the interior of C.

17.31. We define the *area* bounded by a closed curve C to be

$$\int\limits_{C} 1\, dxdy.$$

Since
$$\int\limits_{C} 1\, dxdy = \int\limits_{\alpha}^{\beta}\int\limits_{x_1}^{x_2} dydx = \int\limits_{\alpha}^{\beta} (x_2 - x_1)\, dx$$

this definition leads to the formula for the area bounded by two arcs $y = x_2(x)$, $y = x_1(x)$ and the lines $x = \alpha$, $x = \beta$ which we used in Chapter X.

17.4. Green's theorem

If $P(x, y)$ and $Q(x, y)$ are differentiable inside and on a simple closed curve C, then

$$\int\limits_{C} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \int\limits_{C} P\, dx + \int\limits_{C} Q\, dy.$$

For

$$\int_C \frac{\partial Q}{\partial x}\,dxdy = \int_\gamma^\delta \int_{y_1}^{y_2} \frac{\partial Q}{\partial x}\,dxdy = \int_\gamma^\delta \{Q(y_2, y) - Q(y_1, y)\}\,dy = \int_C Q\,dy$$

and

$$\int_C \frac{\partial P}{\partial y}\,dxdy = \int_\alpha^\beta \int_{x_1}^{x_2} \frac{\partial P}{\partial y}\,dydx = \int_\alpha^\beta \{P(x, x_2) - P(x, x_1)\}\,dx = -\int_C P\,dx,$$

both integrals along C being taken anti-clockwise.

Taking $Q = x$, $P = -y$ we prove another formula for the area A bounded by a closed curve, viz.

$$\int_C (x\,dy - y\,dx) = 2\int_C dxdy = 2A.$$

If C is given by the parametric equations $x = x(t)$, $y = y(t)$, $t_0 \leqslant t \leqslant t_1$, where $x(t)$, $y(t)$ are *differentiable*, then

$$A = \tfrac{1}{2}\int_C (x\,dy - y\,dx) = \tfrac{1}{2}\int_{t_0}^{t_1} (x\dot{y} - \dot{x}y)\,dt,$$

which is the formula introduced in § 10.3.

17.401. $\displaystyle\int_C \{f(x, y) + g(x, y)\}\,dxdy$

$$= \int_C f(x, y)\,dxdy + \int_C g(x, y)\,dxdy.$$

For

$$\int_C (f+g)\,dxdy = \int_\gamma^\delta \left\{\int_{y_1}^{y_2} (f+g)\,dx\right\} dy = \int_\gamma^\delta dy \left\{\int_{y_1}^{y_2} f\,dx + \int_{y_1}^{y_2} g\,dx\right\}$$

$$= \int f\,dxdy + \int g\,dxdy.$$

17.41. If $m \leqslant f(x, y) \leqslant M$ inside and on a simple closed contour C, containing an area A, then

$$mA \leqslant \int_C f(x, y)\,dxdy \leqslant MA.$$

For, by the integral mean-value theorem,

$$\int_{y_1}^{y_2} m\,dx \leqslant \int_{y_1}^{y_2} f(x, y)\,dx \leqslant \int_{y_1}^{y_2} M\,dx$$

and so $$\int_{\gamma}^{\delta}\int_{y_1}^{y_2} m\,dxdy \leqslant \int_{\gamma}^{\delta}\int_{y_1}^{y_2} f(x,y)\,dxdy \leqslant \int_{\gamma}^{\delta}\int_{y_1}^{y_2} M\,dxdy,$$

that is, $$mA \leqslant \int f(x,y)\,dxdy \leqslant MA.$$

17.42. If the interior of a simple closed curve C is divided into two parts contained in simple closed curves C_1 and C_2 then

$$\int_C f(x,y)\,dxdy = \int_{C_1} f(x,y)\,dxdy + \int_{C_2} f(x,y)\,dxdy.$$

Let $f_1(x,y) = f(x,y)$ inside and on C_1 and $f_1(x,y) = 0$ outside C_1, and let $f_2(x,y) = f(x,y)$ inside and on C_2 and $f_2(x,y) = 0$ outside C_2, then $f_1(x,y)+f_2(x,y) = f(x,y)$ throughout C and therefore

$$\int_C f(x,y)\,dxdy = \int_C f_1(x,y)\,dxdy + \int_C f_2(x,y)\,dxdy$$

$$= \int_{C_1} f(x,y)\,dxdy + \int_{C_2} f(x,y)\,dxdy, \quad \text{by 17.3.}$$

17.5. Transformation of a double integral

The transformation $x = X(u, v)$, $y = Y(u, v)$ carries a simple closed curve C into a simple closed curve Γ, and any point inside C into a point inside Γ. The positive value of the Jacobian $J = \dfrac{\partial(X,Y)}{\partial(u,v)}$ and of *at least one* of the derivatives $\partial X/\partial u$, $\partial X/\partial v$, $\partial Y/\partial u$, $\partial Y/\partial v$ exceeds some positive α for all u, v in Γ. Then

$$\int_C f(x,y)\,dxdy = \int_\Gamma f(X,Y)|J|\,dudv.$$

Suppose that $\partial Y/\partial v$ is the derivative which exceeds α, throughout Γ; then we can solve the equation $y = Y(u, v)$ for v, giving, say, $v = \lambda(u, y)$, and therefore $x = X\{u, \lambda(u, y)\}$. Since $|J| \geqslant \alpha > 0$, the equations $x = X(u, v)$, $y = Y(u, v)$ can be solved for u, giving $u = U(x, y)$. We shall show first that the transformation, which leaves y unchanged and which takes x into u by $x = X\{u, \lambda(u, y)\}$, carries C into a simple closed curve C'; this is established by showing that to y correspond two values of u and to u correspond two values of y. Now to y correspond two values of x, y_1, y_2, determined by C, and to y, y_1 corresponds $u^1 = U(y_1, y)$, and to y, y_2 corresponds $u^2 = U(y_2, y)$; thus to y correspond u^1, u^2. To a given

u correspond two values of v, u_1, u_2 by Γ, and so to u correspond $y^1 = X(u, u_1), y^2 = X(u, u_2)$ which completes this part of the proof.

Let $\int_C f(x, y)\, dxdy = \int_\gamma^\delta \int_{v_1}^{v_2} f(x, y)\, dxdy$, and consider $\int_{v_1}^{v_2} f(x, y)\, dx$. Keeping y constant, transform from x to u by means of

$$x = X\{u, \lambda(u, y)\};$$

the integral becomes

$$\int_{u^1}^{u^2} f(X, y)\frac{\partial}{\partial u}X(u, \lambda)\, du$$

and so

$$\int_C f(x, y)\, dxdy = \int_\gamma^\delta \int_{u^1}^{u^2} f(X, y)\frac{\partial X}{\partial u}\, dudy$$

$$= \theta \int f(X, y)\frac{\partial X}{\partial u}\, dudy,$$

where $\theta = +1$ or -1 according as $u^2 > u^1$ or $u^2 < u^1$.

Now, with appropriate limits,

$$\int_{C'} f(X, y)\frac{\partial X}{\partial u}\, dudy = \int\int f(X, y)\frac{\partial X}{\partial u}\, dydu.$$

Consider $\int f(X, y)\frac{\partial X}{\partial u}\, dy$; transform y into v, keeping u constant, by means of $y = Y(u, v)$. The integral becomes

$$\int f\frac{\partial}{\partial u}X(u, \lambda)\frac{\partial Y}{\partial v}\, dv,$$

with appropriate limits.

Hence

$$\int_{\gamma'} f(X, y)\frac{\partial X}{\partial x}\, dudy = \int\int f\frac{\partial}{\partial u}X(u, \lambda)\frac{\partial Y}{\partial v}\, dvdu$$

$$= \int f(X, Y)\frac{\partial}{\partial u}X(u, \lambda)\frac{\partial Y}{\partial v}\, dudv,$$

apart from an ambiguity in sign.

Thus

$$\int_C f(x, y)\, dxdy = k\int_\Gamma f(X, Y)\frac{\partial}{\partial u}X(u, \lambda)\frac{\partial Y}{\partial v}\, dudv,$$

where $k = +1$ or $k = -1$.

But $$\frac{\partial Y}{\partial v}\frac{\partial}{\partial u}X(u,\lambda)=\frac{\partial Y}{\partial v}\frac{\partial X}{\partial u}+\frac{\partial Y}{\partial v}\frac{\partial \lambda}{\partial u}\frac{\partial X}{\partial v},$$

and since $y = Y\{u, \lambda(u, y)\}$ identically,

$$0=\frac{\partial Y}{\partial u}+\frac{\partial Y}{\partial v}\frac{\partial \lambda}{\partial u},$$

whence $$\frac{\partial Y}{\partial v}\frac{\partial}{\partial u}X(u,\lambda)=\frac{\partial X}{\partial u}\frac{\partial Y}{\partial v}-\frac{\partial X}{\partial v}\frac{\partial Y}{\partial u}=J,$$

and therefore

$$\int_C f(x,y)\,dxdy = k\int_\Gamma f(X,Y)J\,dudv.$$

The value of k depends only upon the transformation functions X, Y and not upon $f(x, y)$, and may therefore be determined by giving $f(x, y)$ some special value. Take $f(x, y) = 1$, then

$$\int_C dxdy = \int_\Gamma kJ\,dudv.$$

Since $\int_C dxdy$ represents the area bounded by C it is positive; furthermore, as J is continuous and non-zero, it is of constant sign. Hence $\int_\Gamma kJ\,dudv$ has the sign of kJ and so kJ is positive; but $k = \pm 1$, and so $kJ = |J|$, whence

$$\int_C f(x,y)\,dxdy = \int_\Gamma f(X,Y)|J|\,dudv.$$

We may dispense with the condition that the positive value of one of the first derivatives of X or Y exceeds α throughout Γ, provided that we know instead that the transformation $x = X(u, v)$, $y = Y(u, v)$ is such that any division of the interior of C by simple closed curves is transformed into a division of the interior of Γ by simple closed curves; for the integral over C is the sum of the integrals over the dividing curves, and so too the integral over Γ is the sum of the integrals over the corresponding curves, and (as we saw in 15.83) if Γ is divided into sufficiently small parts, in each part the positive value of one of the derivatives of X or Y exceeds a definite positive number throughout that part (since $|J| \geqslant \alpha$). Hence we may apply the foregoing proof to each simple closed curve by which Γ is subdivided, and, by addition, the theorem is proved for the whole curve Γ.

An example of a transformation of this kind is

$$x = u\cos v, \qquad y = u\sin v.$$

To the line $u = p$ corresponds the circle $x^2+y^2 = p^2$, to the line $v = \lambda$ corresponds the half-line $x/\cos\lambda = y/\sin\lambda$, and therefore to a subdivision of a rectangle $u = p$, $u = q$, $v = \lambda$, $v = \mu$ into any number of rectangular parts corresponds a subdivision of the ring sector, formed by the circles $x^2+y^2 = p^2$, $x^2+y^2 = q^2$, and the radii $x/\cos\lambda = y/\sin\lambda$, $x/\cos\mu = y/\sin\mu$ into parts which are also ring sectors.

EXAMPLE. To prove $\int\limits_0^\infty e^{-x^2}\,dx = \frac{1}{2}\sqrt{\pi}$.

We observe first that

$$\int\limits_0^R\int\limits_0^R e^{-(x^2+y^2)}\,dxdy = \int\limits_0^R e^{-y^2}\left\{\int\limits_0^R e^{-x^2}\,dx\right\}dy = \left(\int\limits_0^R e^{-x^2}\,dx\right)^2.$$

Divide the square $(0, R)(0, R)$ into three simple closed curves; C_1 formed by the quadrant of $x^2+y^2 = \epsilon^2$ bounded by the positive x- and y-axes; C_2 formed by the quadrants of $x^2+y^2 = \epsilon^2$, $x^2+y^2 = R^2$ and the positive axes from ϵ to R; C_3 formed by the quadrant of $x^2+y^2 = R^2$ and the lines $x = R$, $y = R$ from $y = 0$ to $y = R$, and from $x = 0$ to $x = R$ respectively.

In C_1, $e^{-(x^2+y^2)} \leqslant 1$ and the area bounded by C_1 is $\frac{1}{4}\pi\epsilon^2$, and so, by 17.41, $\int\limits_{C_1} e^{-x^2-y^2}\,dxdy \leqslant \frac{1}{4}\pi\epsilon^2$. In C_3, $x^2+y^2 \geqslant R^2$ and the area bounded is $R^2(1-\frac{1}{4}\pi)$ and so $\int\limits_{C_3} e^{-x^2-y^2}\,dxdy \leqslant (1-\frac{1}{4}\pi)R^2e^{-R^2}$.

Lastly, we consider $\displaystyle\int\limits_{C_2} e^{-(x^2+y^2)}\,dxdy.$

Under the transformation $x = u\cos v$, $y = u\sin v$ the upper half of the circle $x^2+y^2 = \epsilon^2$ becomes the line $u = \epsilon$, the upper half of $x^2+y^2 = R^2$ becomes $u = R$, the axis $y = 0$ becomes $v = 0$, and the axis $x = 0$ becomes $v = \frac{1}{2}\pi$, and therefore C_2 transforms into a rectangle. The Jacobian of the transformation is u, which is never less than ϵ in the transform of C_2. Hence

$$\int\limits_{C_2} e^{-(x^2+y^2)}\,dxdy = \int\limits_0^{\frac{1}{2}\pi}\int\limits_\epsilon^R e^{-u^2}u\,dudv = \tfrac{1}{4}\pi\{e^{-\epsilon^2}-e^{-R^2}\}.$$

Since

$$\int_0^R \int_0^R e^{-(x^2+y^2)}\,dxdy = \int_{C_1} + \int_{C_2} + \int_{C_3} e^{-(x^2+y^2)}\,dxdy \quad \text{by 17.42,}$$

therefore

$$\left|\left(\int_0^R e^{-x^2}\,dx\right)^2 - \tfrac{1}{4}\pi\right| \leqslant \tfrac{1}{4}\pi\{e^{-\epsilon^2}-1-e^{-R^2}\}+\tfrac{1}{4}\pi\epsilon^2+(1-\tfrac{1}{4}\pi)R^2e^{-R^2}.$$

This inequality holds for all positive ϵ and R. Let $\epsilon \to 0$ and $R \to \infty$, then $\left(\int_0^\infty e^{-x^2}\,dx\right)^2 = \frac{1}{4}\pi$, which completes the proof.

APPENDIX

THE ARITHMETIC OF ENDLESS DECIMALS. FORMAL PROOF OF THE EXISTENCE OF THE LIMIT OF A CONVERGENT SEQUENCE OF DECIMALS

In the first chapter we described in general terms the way in which the limit of a convergent sequence is constructed, digit by digit, but in that description we omitted the consideration of certain difficulties which may arise in the construction process; to meet these difficulties we shall now give a formal proof of the existence of the limit.

.1. Whether x be a terminating or endless decimal we define $(x)_k$ to be the value of x to k decimal places, i.e. if $x = \pm a_0 \cdot a_1 a_2 a_3 \ldots$, then $(x)_k = \pm a_0 \cdot a_1 a_2 \ldots a_k$. For instance $(17\cdot 2\dot{6})_4 = 17\cdot 2666$ and $(-\sqrt{2})_2 = -1\cdot 41$. It follows that $10^k (x)_k$ is a whole number, for any x and k, and $|(x)_k| = |x|_k$, writing $|x|_k$ for $(|x|)_k$.

.11. If x is a positive terminating decimal, then

$$0 \leqslant x-(x)_k < 1/10^k$$

for any k; for if $x = a_0 \cdot a_1 a_2 \ldots a_k a_{k+1} \ldots a_p$ (where a_r may be zero from any stage onwards) then $x-(x)_k = \cdot 00\ldots 0\, a_{k+1} \ldots a_p$ which is non-negative and less than $1/10^k$.

.12. If x and y are positive terminating decimals such that, for some r, $(x-y)_r = 0$ then $(x)_r = (y)_r + \theta/10^r$, where θ has one of the values -1, 0, $+1$.

Proof. Since $|x-y| < 1/10^r$ therefore $-1 < (x-y)10^r < 1$; furthermore

$$0 \leqslant 10^r\{x-(x)_r\} < 1 \quad \text{and} \quad 0 \leqslant 10^r\{y-(y)_r\} < 1$$

and so

$$\begin{aligned} 10^r\{(x)_r-(y)_r\} &= 10^r\{y-(y)_r-[x-(x)_r]+x-y\} \\ &\leqslant 10^r\{y-(y)_r+x-y\} \\ &< 2 \end{aligned}$$

and

$$\begin{aligned} 10^r\{(x)_r-(y)_r\} &\geqslant 10^r\{-[x-(x)_r]+x-y\} \\ &> -2. \end{aligned}$$

Denote $10^r\{(x)_r-(y)_r\}$ by θ, then θ is a whole number between -2 and $+2$ exclusive, proving that θ takes one of the values $-1, 0, +1$.

.121. If x is any decimal and if $p > n$ then $(x)_p-(x)_n = 0(n)$; for if $x = \pm a_0{\cdot}a_1a_2a_3...$, then

$$(x)_p-(x)_n = \pm(a_0{\cdot}a_1a_2...a_p-a_0{\cdot}a_1a_2...a_n)$$
$$= \pm(0{\cdot}00...0\,a_{n+1}...a_p) = 0(n).$$

.13. If x is a positive decimal, terminating or endless, and if, for $p > r$,

$$(x)_p-(x)_r < 1/10^p,$$

then $(x)_p = (x)_r$.

For $(x)_p \geqslant (x)_r$, and so $10^p(x)_p-10^p(x)_r$ is a non-negative integer *less* than unity, which proves that $10^p(x)_p-10^p(x)_r = 0$.

The limit of a sequence of terminating decimals

.2. If with a sequence of terminating decimals (a_n) we can associate a sequence of integers ν_r and a terminating or endless decimal λ such that, for every k and $n \geqslant \nu_k$, $a_n = (\lambda)_n+0(k)$ then λ is said to be the limit of the sequence (a_n), and we write $a_n \to \lambda$.

.21. If $a_n \to \lambda$ then $a_n-(\lambda)_n \to 0$, for the one equation,

$$a_n-(\lambda)_n = 0(k),$$

expresses both of these facts.

.3. If $(\mathbf{a_n})$ is a convergent sequence of positive terminating decimals then

either (1) for all r, $(\mathbf{a_n})_\mathbf{r} = (\mathbf{a_{p_r}})_\mathbf{r}$, $\mathbf{n \geqslant p_r}$,

or (2) for a certain k, $\mathbf{a_n \to (a_{\nu_k})_k}$,

or (3) for a certain k, $\mathbf{a_n \to (a_{\nu_k})_k+1/10^k}$.

Proof. Since (a_n) is convergent we can determine ν_r such that, for all r and $n \geqslant \nu_r$,

$$(a_n-a_{\nu_r})_r = 0.$$

Hence by .12

$$(a_n)_r = (a_{\nu_r})_r+\theta_n^r/10^r, \quad \text{each } \theta_n^r = -1,\ 0, \text{ or } +1.$$

There are the following possibilities to be considered:

(*a*) For all r and for n not less than a certain p_r (where $p_r \geqslant \nu_r$), $\theta_n^r = 0$; then for all r

$$(a_n)_r = (a_{p_r})_r, \quad n \geqslant p_r.$$

(*b*) For all r and for $n \geqslant p_r$, $\theta_n^r = 1$; then for all r

$$(a_n)_r = (a_{p_r})_r, \quad n \geqslant p_r.$$

(c) For all r and for $n \geqslant p_r$, $\theta_n^r = -1$; then for all r

$$(a_n)_r = (a_{p_r})_r, \quad n \geqslant p_r.$$

(d) For a fixed r, θ_n^r takes both the values $+1$ and -1 for arbitrarily great values of n. *We shall prove this cannot happen.*

For suppose that for a certain r, θ_n^r takes the value $+1$ for arbitrarily great values of n and let j be such a value of n.

Then $$(a_j)_r = (a_{\nu_r})_r + 1/10^r.$$

Choose $j \geqslant \nu_{r+1}$ and $k > j$, so that $(a_k)_r = (a_{\nu_r})_r + \theta_k^r/10^r \leqslant (a_j)_r$.

Since
$$-1/10^{r+1} < a_j - a_{\nu_{r+1}} < 1/10^{r+1},$$
$$-1/10^{r+1} < a_k - a_{\nu_{r+1}} < 1/10^{r+1}$$

and therefore $$|a_j - a_k| < 2/10^{r+1} < 1/10^r,$$

whence $$0 \leqslant (a_j)_r - (a_k)_r < a_j - (a_k - 1/10^r) < 2/10^r.$$

But $$(a_k)_r = (a_{\nu_r})_r + \theta_k^r/10^r$$

and so $1 - \theta_k^r < 2$, i.e. $\theta_k^r > -1$, so that $\theta_k^r = 0$ or 1. Thus θ_n^r cannot take both the values $+1$, -1 for arbitrarily great values of n.

(e) For a certain r, θ_n^r takes the values 0, 1 and *only these values* for arbitrarily great values of n. Then $(a_j)_r = (a_{\nu_r})_r + 1/10^r$ and $(a_k)_r = (a_{\nu_r})_r$ for j and k as great as we please, and so

$$(a_j)_r = (a_k)_r + 1/10^r.$$

But if j, k are sufficiently great (and $p > r$) then $(a_j - a_k)_p = 0$ and so
$$(a_j)_p = (a_k)_p + \phi/10^p, \qquad \phi = -1,\ 0, \text{ or } 1.$$

Therefore
$$\begin{aligned} 0 \leqslant (a_j)_p - (a_j)_r &= (a_k)_p - (a_k)_r + \phi/10^p - 1/10^r \\ &< 1/10^r + \phi/10^p - 1/10^r \\ &= \phi/10^p \leqslant 1/10^p. \end{aligned}$$

Hence, by .13, $(a_j)_p = (a_j)_r = (a_{\nu_r})_r + 1/10^r = l$ (say). Since every sufficiently great value of n is either a j or a k, therefore

$$(a_n)_p = l \quad \text{or} \quad (a_n)_p = l - \phi/10^p$$

and so $$|l - a_n| = |l - (a_n)_p + (a_n)_p - a_n| < 2/10^p,$$

which proves that
$$a_n \to l = (a_{\nu_r})_r + 1/10^r.$$

(*f*) For a certain r, θ_n^r takes the values -1, 0 and *only these values* for all sufficiently great values of n. Then, exactly as in (*e*) above, we can prove that $a_n \to (a_{\nu_r})_r$.

Since the function θ_n^r is either a constant function of n, *for all values of* r, or there is a value of r for which θ_n^r varies with n, the cases listed in (*a*), (*b*), (*c*), (*d*), (*e*), and (*f*) (which are mutually exclusive) exhaust all the possibilities, which proves that one (and only one) of the conditions (1), (2), (3) above is necessarily true for any positive convergent sequence.

.31. If the sequence (a_n) is such that, for all r and $n \geqslant \nu_r$,

$$(a_n)_r = (a_{\nu_r})_r$$

and if $l_0, l_1, l_2, \ldots, l_r$ are the whole part and the first r decimal figures of a_{ν_r} then the endless decimal $\lambda = l_0 \cdot l_1 l_2 l_3 \ldots$ is the limit of the sequence (a_n).

For

$$\begin{aligned} a_r - (\lambda)_r &= a_r - (a_{\nu_r})_r \\ &= \{a_n - (a_n)_r\} + \{a_r - (a_r)_p\} + \{(a_n)_p - a_n\} + \{(a_r)_p - (a_n)_p\} \\ &\leqslant 1/10^r + 1/10^p + 1/10^p, \quad \text{if } r \geqslant \nu_p,\ n \geqslant \nu_r \geqslant \nu_p, \\ &< 1/10^{p-1} \end{aligned}$$

and so $a_r - (\lambda)_r \to 0$, proving that $a_r \to \lambda$.

.32. From .3 and .31 it follows that if (a_n) is a convergent sequence of positive terminating decimals then there is determined a decimal λ, terminating or otherwise, which is the limit of the sequence (a_n).

.33. If (a_n) is convergent and a_n takes both positive and negative values for arbitrarily great values of n, then $a_n \to 0$.

Choose r, s so that a_r is positive, a_s negative; then if r and s are great enough

$$a_r - a_s = 0(k)$$

and therefore

$$0 \leqslant a_r \leqslant a_r - a_s = 0(k) \quad \text{and} \quad 0 \leqslant -a_s \leqslant a_r - a_s = 0(k).$$

Thus $a_r = 0(k)$ and $a_s = 0(k)$, proving $a_n \to 0$.

.34. If (a_n) is a convergent sequence of terminating decimals, positive or negative, then there is a decimal λ which is the limit of (a_n).

For

either (1) all $a_n \geqslant 0$ from some n onwards,

or (2) all $a_n \leqslant 0$ from some n onwards,

or (3) a_n takes both positive and negative values for arbitrarily great values of n.

In case (1) the limit is given by .32. In case (2) $-a_n \geqslant 0$ and so, by .32, $-a_n \to \lambda$, whence $a_n \to -\lambda$. In case (3) the limit is zero by .33.

.4. If α and β are endless decimals, such that $(\alpha)_r-(\beta)_r \to 0$, then we say that α and β are equal and write $\alpha = \beta$.

To justify this generalization of the notion of equality we must show that when α and β are terminating decimals and $(\alpha)_r-(\beta)_r \to 0$ then $\alpha = \beta$ can be proved true For if k is a number greater than the number of decimal figures in either of the terminating decimals α and β then, for $r \geqslant k$, $(\alpha)_r = \alpha$, $(\beta)_r = \beta$, and $(\alpha-\beta)_r = \alpha-\beta$, and therefore, since $(\alpha)_r-(\beta)_r = 0(k)$ provided $r \geqslant$ some N_k, it follows that $\alpha-\beta = (\alpha-\beta)_k = \{(\alpha)_r-(\beta)_r\}_k = 0$, i e. $\alpha = \beta$.

.401. If α and β are endless decimals then the sequences

$$(\alpha)_n+(\beta)_n, \quad (\alpha)_n-(\beta)_n, \quad \text{and} \quad (\alpha)_n(\beta)_n$$

are convergent; hence $(\alpha)_n+(\beta)_n$ tends to a limit and we call this limit the sum of α and β, denoted by $\alpha+\beta$. Similarly $\alpha-\beta$ and $\alpha\beta$ denote the limits of $(\alpha)_n-(\beta)_n$, $(\alpha)_n(\beta)_n$ and are called the difference and product of α and β respectively.

.402. Since $\alpha+\beta$ and $\alpha-\beta$ are the limits of the sequences $(\alpha)_n+(\beta)_n$ and $(\alpha)_n-(\beta)_n$, therefore (by .21)

$$(\alpha+\beta)_n-\{(\alpha)_n+(\beta)_n\} \to 0$$

and

$$(\alpha-\beta)_n-\{(\alpha)_n-(\beta)_n\} \to 0.$$

Accordingly $\alpha = \beta$ if, for every r, $(\alpha-\beta)_r = 0$, for $(\alpha-\beta)_r = 0$ involves $(\alpha)_r-(\beta)_r \to 0$.

.403. If $(a+b) = c$ then $a = (c-b)$; for $(a+b)_r-(c)_r \to 0$ and so, by .402, $(a)_r+(b)_r-(c)_r \to 0$, i.e. $(a)_r-\{(c)_r-(b)_r\} \to 0$, whence $(a)_r-(c-b)_r \to 0$, proving $a = (c-b)$.

.404. For any decimals a, b, c we have $a(b+c) = ab+ac$. Since $[a(b+c)]_n-(a)_n(b+c)_n \to 0$ and $(b+c)_n-\{(b)_n+(c)_n\} \to 0$ therefore $[a(b+c)]_n-(a)_n\{(b)_n+(c)_n\} \to 0$, and similarly

$$(ab+ac)_n-(a)_n\{(b)_n+(c)_n\} \to 0,$$

whence

$$[a(b+c)]_n-[ab+ac]_n \to 0,$$

which proves $a(b+c) = ab+ac$.

.41. If α is an endless decimal and $(\alpha)_k = 0$, we write $\alpha = 0(k)$.

.42. A sequence of endless decimals (a_n) is said to be convergent if $a_p-a_n = 0(k)$ for $p \geqslant n \geqslant N_k$.

.43. If α is an endless decimal then $\alpha-(\alpha)_k = 0(k)$.

For

$$(\alpha)_n-\{(\alpha)_k\}_n = (\alpha)_n-(\alpha)_k, \quad \text{if } n \geqslant k,$$
$$= (\alpha)_n-\{(\alpha)_n\}_k;$$

let $\alpha = p_0{\cdot}p_1p_2p_3\ldots$, then $(\alpha)_n = p_0{\cdot}p_1p_2\ldots p_n$, and so

$$(\alpha)_n-\{(\alpha)_k\}_n = {\cdot}00\ldots0\,p_{k+1}p_{k+2}\ldots p_n \to {\cdot}00\ldots0\,p_{k+1}p_{k+2}\ldots \text{ (by .31)}$$

but $(\alpha)_n-\{(\alpha)_k\}_n \to \alpha-(\alpha)_k$ and therefore

$$\alpha-(\alpha)_k = {\cdot}00\ldots0\,p_{k+1}p_{k+2}\ldots, \quad \text{i.e. } \{\alpha-(\alpha)_k\}_k = 0,$$

so that, by .41,

$$\alpha-(\alpha)_k = 0(k).$$

The limit of a sequence of endless decimals

.44. If (a_n) is a sequence of endless decimals, and if λ is a decimal, terminating or endless, such that $\lambda-a_n = 0(k)$ for $n \geqslant N_k$, then λ is called the *limit* of (a_n) and we write $a_n \to \lambda$. If a_n is a terminating decimal this definition is consistent with .21, for

$$a_n-(\lambda)_n = a_n-\lambda+\lambda-(\lambda)_n = 0(k)+0(n) = 0(k-1),$$

if $n \geqslant \max(k, N_k)$.

.45. The limit of a sequence (a_n) is *unique*, for if λ and λ' satisfy $\lambda-a_n = 0(k)$, $\lambda'-(a_n) = 0(k)$ then $\lambda-\lambda' = 0(k-1)$, for any k, so that, by .402, $\lambda = \lambda'$.

.5. If (a_n) is a convergent sequence of endless decimals then the sequence of terminating decimals $(a_p)_p$ is convergent.

For

$$(a_p)_p-(a_n)_n = \{(a_p)_p-a_p\}+\{a_p-a_n\}+\{a_n-(a_n)_n\}$$
$$= 0(p)+0(k)+0(n), \quad \text{provided } p \geqslant n \geqslant N_k;$$
$$= 0(k-1), \quad \text{provided } p \geqslant n \geqslant \max(k, N_k).$$

.51. If (a_n) is any convergent sequence then we can determine a decimal λ which is the limit of the sequence.

By .5, $(a_p)_p$ is convergent, and so, by .34, we can determine its limit λ. Then

$$\begin{aligned}\lambda-a_n &= \lambda-(a_n)_n+(a_n)_n-a_n \\ &= 0(k)+0(n), \quad \text{if } n \geqslant N_k, \\ &= 0(k-1), \quad \text{if } n \geqslant \max(k, N_k),\end{aligned}$$

proving that λ is the limit of the sequence (a_n).

Theorem .51 contains Theorems .32 and .34 as special cases, but .32 and .34 are essential steps in preparing for .51. The arithmetic of endless decimals is based upon convergent sequences of terminating decimals, the properties of which must therefore be established before we can even formulate Theorem .51. Whereas the proof of Theorem .34 involves an extension of the number concept, a convergent sequence of *terminating* decimals having an *endless* decimal for limit, no further extension is involved in Theorem .51, the limit of a convergent sequence of endless decimals being also an endless decimal. Endless decimals have therefore a self-sufficiency which terminating decimals lack, as far as operations with convergent sequences are concerned.

.6. We *define* $1/\beta$ to be the limit of the *convergent* sequence $1/(\beta)_n$, provided that for some μ, $|\beta|_\mu > 0$. (Observe that we cannot include this definition in .401, since $1/(\beta)_n$ is an endless decimal, the quotient of the integer 10^n divided by the integer $10^n(\beta)_n$. If $[x/y]$ denotes the greatest number of times y is contained in x, then the formal definition of m/n, for non-zero integers m, n, is

$$(m/n)_p = [10^p m/n]/10^p.$$

Since $y[x/y] \leqslant x < y[x/y]+y$, therefore

$$n.(m/n)_p \leqslant m < n\{(m/n)_p+10^{-p}\},$$

i.e. $$m-n/10^p < n.(m/n)_p \leqslant m,$$

which proves that $n.(m/n)_p \to m$; but, for any p, $(n)_p = n$ and therefore $(n)_p.(m/n)_p \to m$, i.e., by .401, $n.(m/n) = m$. Similarly, if a, b, c, d are integers, $(a/b).(c/d) = ac/bd$, etc.)

.61. If, for some μ, $|\beta|_\mu > 0$, then $\beta \times (1/\beta) = 1$.

Since $1/\beta$ is the limit of $1/(\beta)_n$ it follows that $(1/\beta)_n - 1/(\beta)_n \to 0$ and therefore $(\beta)_n(1/\beta)_n - (\beta)_n . 1/(\beta)_n \to 0$, whence

$$(\beta)_n(1/\beta)_n - 1 \to 0,$$

which proves that $\beta(1/\beta) = 1$.

.62. We define α/β to equal $\alpha(1/\beta)$.

EXAMPLES

I

1. Prove that for any decimals x, y, $y \neq 0$,

$$|1+x| \leqslant 1+|x|, \qquad |xy| = |x||y|, \qquad |x|/|y| = |x/y|.$$

Deduce that $|x+y| \leqslant |x|+|y|$, $|x-y| \geqslant ||x|-|y||$.

1.01. If $x > -|x|$ prove that $x > 0$.

1.1. Prove that for any decimals $a_1, a_2, a_3, \ldots, a_n$,

$$\left|\sum_1^n a_r\right| \leqslant \sum_1^n |a_r|.$$

1.11. If $s_n \to l$ and $1/p_n \to 0$, prove that $s_{p_n} \to l$, p_n being an integer for all n.

1.2. Show that the positive series $\sum a_n$ is convergent if $(a_n)^{1/n}$ tends to a limit which is less than unity. Hence show that $\sum x^n/n^n$ is convergent for all values of x.

1.201. If $s_n \to l$ prove that $|s_n| \to |l|$.

1.21. Prove that $\sum 1/n$ diverges, and that by taking a sufficient number of terms the sum exceeds any chosen number.

1.211. Test for convergence the series $\sum n/(n^2+1)$, $\sum x^n/n(n+1)$.

1.22. Find the sum to n terms of the series $\sum\limits_{n=1} n!/(x+1)(x+2)\ldots(x+n)$ and hence show that the series converges for $x > 1$ and diverges for $x \leqslant 1$.

1.221. If $a_n/b_n \to l > 0$ then $\sum a_n$ and $\sum b_n$ converge and diverge together.

1.23. Prove the convergence of

$$\sum 1/n(n+1), \qquad \sum 1/n(n+1)(n+2), \qquad \sum 1/n(n+1)\ldots(n+r-1), \qquad \sum 1/n^r,$$

where r is an integer greater than unity.

1.3. If a_n is positive and decreases steadily to zero, prove that $\sum(-1)^n a_n$ converges.

1.31. If a_n is positive and $a_n/a_{n+1} \geqslant 1+k/n$, $k > 0$, $n \geqslant N$, prove that $\sum(-1)^n a_n$ converges.

1.32. If a_n is positive and decreases steadily, and if $\sum a_n$ is convergent, prove that $na_n \to 0$.

1.4. Prove that $k+4$ terms of the series $\sum\limits_{r=0} 1/r!$ suffice to determine the limit of the series to k places of decimals; determine the limit to 3 places.

1.401. Prove that $\sum (n!)^2/(2n)!$ and $\sum(-1)^{n+1}/n^4$ converge, and evaluate the second limit to 2 decimal places.

1.402. Discuss the convergence of the series $\sum 2^n x^n/n^{\frac{1}{2}}$.

1.41. If a series is absolutely convergent, prove that no rearrangement of its terms can alter its sum.

1.5. The product of any number of endless decimals is independent of the order of its factors.

1.501. Assuming the index laws

$$p^m . p^n = p^{m+n}, \qquad (p^m)^n = p^{mn}, \qquad (pq)^m = p^m q^m$$

for integers p, q, m, n and defining $(p/q)^m = p^m/q^m$, prove that

$$(p/q)^m \cdot (p/q)^n = (p/q)^{m+n}, \quad \{(p/q)^m\}^n = (p/q)^{mn}, \quad \{(p/q)(r/s)\}^m = (p/q)^m(r/s)^m.$$

1.51. If x is a positive endless decimal, x_k its value to k decimal places (x_0 being the whole part of x), and if p is a positive integer, prove that the sequence $x_0^p, x_1^p, x_2^p,\ldots$ is convergent. The limit of this sequence is denoted by x^p.

1.511. If x, y are positive endless decimals and m, n are positive integers prove that

$$x^m . x^n = x^{m+n}, \qquad (x^m)^n = x^{mn}, \qquad (xy)^m = x^m . y^m.$$

1.6. If $\varphi(n)$ is a positive strictly decreasing function then $\sum \varphi(n)$ converges or diverges together with $\sum 2^n \varphi(2^n)$.

1.61. Prove that $\sum 1/n^\sigma$ converges if $\sigma > 1$ and diverges if $\sigma \leqslant 1$.

1.62. If $u_n/u_{n+1} \geqslant v_n/v_{n+1} > 0$, $n \geqslant m$, then $\sum u_n$ converges if $\sum v_n$ converges, and $\sum v_n$ diverges if $\sum u_n$ diverges.

1.63. If $u_n/u_{n+1} = 1+\beta/n+\theta_n/n^{1+\lambda}$, $|\theta_n| < M$, $\lambda > 0$, prove that if $\beta > 1$ then $nu_n \to 0$ and $\sum u_n$ converges, and if $\beta < 1$ then $\sum u_n$ diverges.

1.64. If $|a_n/a_{n+1}| \to R$ prove that the three series $\sum a_n x^n$, $\sum na_n x^{n-1}$, and $\sum a_n \dfrac{x^{n+1}}{n+1}$ have the same interval of convergence.

1.7. If the interval (a_{n+1}, b_{n+1}) is contained in the interval (a_n, b_n) for all n, and if $b_{n+1}-a_{n+1} \leqslant k(b_n-a_n)$, $0 < k < 1$, then a_n and b_n tend to a common limit which lies in every interval (a_n, b_n).

✓1.71. If $a_{n+1} = \frac{1}{2}(a_n+b_n)$, $b_{n+1} = +\sqrt{(a_n b_n)}$, $0 < b_0 < a_0$, prove that a_n and b_n tend to a common limit.

1.72. If $a_{n+1} = +\sqrt{(a_n+k)}$, $a_0 > \frac{1}{2}$, $k > 0$, prove that the sequence a_n is convergent and that its limit is the positive root of the equation

$$x^2-x-k = 0.$$

1.73. If $a_{n+1} = k/(1+a_n)$, $k > 0$, $a_0 > 0$, prove that a_n converges to the positive root of the equation $x^2+x-k = 0$.

1.74. If $x_{n+1} = 2a^2x_n/(x_n^2+a^2)$ and $y_{n+1} = (y_n^2+a^2)/2y_n$, $x_0 > 0$, $y_0 > 0$, $a > 0$, then both x_n and y_n tend to a.

✓1.75. If $x_{n+1} = \frac{1}{2}(x_n+x_{n-1})$ prove that $x_n \to (x_0+2x_1)/3$.

1.76. If $|b_{n+2}-b_{n+1}| \leqslant k|b_{n+1}-b_n|$ and $k < 1$, prove that the sequence b_n is convergent. Hence show that, if for all n,

$$pc_{n+2}-(p+q)c_{n+1}+qc_n = 0, \quad p > q > 0,$$

then $pc_{n+2}-qc_{n+1} = pc_1-qc_0$ and c_n converges to $(pc_1-qc_0)/(p-q)$.

1.77. If $y = (\alpha x+\beta)/(x+\gamma)$ and if λ, μ are the roots of the equation

$$x^2-(\alpha-\gamma)x-\beta = 0$$

show that
$$\frac{y-\lambda}{y-\mu} = \frac{\gamma+\mu}{\gamma+\lambda}\frac{x-\lambda}{x-\mu}.$$

Hence prove that if $a_n a_{n+1}-p(a_n-a_{n+1}) = q^2$, $n \geqslant 0$, then

$$a_n \to q \quad \text{if} \quad p/q > 0, \quad \text{and} \quad a_n \to -q \quad \text{if} \quad p/q < 0.$$

1.8. If $a_{n+2}+pa_{n+1}+qa_n = 0$, $n > 0$, prove that $\sum a_n x^n$ converges for all sufficiently small values of $|x|$ and that the limit of the series is

$$\{a_0+(a_1+pa_0)x\}/(1+px+qx^2).$$

1.9. If $f(x)$ is a strictly increasing (or strictly decreasing) function and $f(n) \to l$ and if $g(n) > 0$ and $1/g(n) \to 0$, prove that $f(g(n)) \to l$. Illustrate the necessity for the condition '$f(x)$ is increasing or $f(x)$ is decreasing' by an example.

II

2. Prove that $\sqrt{x}$ is continuous in the interval $(0, c)$ for any $c > 0$.

2.1. If $q > p^2$ prove that $\sqrt{(x^2+2px+q)}$ is continuous in any interval.

2.2. Prove that $a_0+a_1x+a_2x^2+...+a_nx^n$ is continuous in any interval.

2.3. If $a_0+a_1x+a_2x^2+...+a_nx^n = b_0+b_1x+b_2x^2+...+b_nx^n$ for all x *near* $x = 0$ prove that the two polynomials are equal for *all* values of x.

2.4. At each point x of an interval (a, b) at least one of the continuous functions $f(x)$, $g(x)$ takes a value greater than a certain positive number δ. Prove that (a, b) can be divided into a finite number of parts such that in each part i_k, *either* $f(x) > \frac{1}{2}\delta$ for all x in i_k *or* $g(x) > \frac{1}{2}\delta$ for all x in i_k.

2.5. The function $g(x)$ is continuous in $[a, b]$ and $g(x)$ is not zero for any x in $[a, b]$. Prove that $g(x)$ is of constant sign in $[a, b]$.

2.6. If $f(x)$ is continuous in (a, b) and $f(x) = \lambda$ for all *terminating* decimals in (a, b), prove that $f(x) = \lambda$ at any point in (a, b).

Deduce that if $f(x)$ and $g(x)$ are continuous in (a, b) and $f(x)$ $g(x)$ for all terminating decimals in (a, b) then the equality holds at all points of (a, b).

2.7. $a_{n+1} = k/(1+a_n)$, $a_0 > 0$, $k > 0$; prove that, for any n, a root of $x^2+x = k$ lies between a_{2n} and a_{2n+1}, provided $a_0^2+a_0 \neq k$.

2.8. A function f(x) is said to be *semi-continuous* in an interval (a, b], if given any k we can determine points $a = a_0^k, a_1^k, a_2^k, ..., a_{\nu_k}^k = b$ such that $f(X)-f(x) = 0(k)$ for any x, X in the same *half-open* sub-interval $(a_r^k, a_{r+1}^k]$. The points a_r^k, $r = 0, 1, ..., \nu_k$ are said to form a k-chain of f(x) in (a, b].

Prove that if $f(x)$ is semi-continuous in $(a, b]$ then $|f(x)|$ is semi-continuous in $(a, b]$, and also that $f(x)$ is bounded in $(a, b]$.

2.81. If $f(x)$ is semi-continuous in $(a, b]$, then show that $f(x)$ is semi-continuous in any part of $(a, b]$.

2.82. Prove that a continuous function is semi-continuous.

2.83. If $f(x)$ and $g(x)$ are semi-continuous in $(a, b]$ prove that $f(x)+g(x)$ and $f(x)g(x)$ are semi-continuous in $(a, b]$.

III

3. Write down the derivatives of the following functions:

$x(1+x)$, $(1+x^2)^{\frac{1}{2}}$, $(1-x)^{\frac{1}{2}}(1+x)^{\frac{3}{2}}$, $(a+bx)^2$, $(x+a)/(x+b)$,
$(x^4+1)/(x^2+1)^2$, $(2x^2-1)/(2x^2+3x+1)$, $(x+a)(x+b)^2(x+c)^3$,
$(2x+a+b)/\{x^2+(a+b)x+ab\}$, $1/\{\sqrt{(x^2+1)}-x\}$, $(1+\sqrt{x})/(1-\sqrt{x})$,
$1/\{\sqrt{(x+1)}-\sqrt{(x-1)}\}$, $1/\{\sqrt{(1+x)}-\sqrt{(1-x)}\}$, $(1+x+x^2)^{\frac{1}{2}}/(1-x+x^2)^{\frac{1}{2}}$,
$\{x^{\frac{1}{2}}+\ {}^{\frac{1}{2}}\}^2$, $(1+x)/(a+bx)^2$, $(a+bx)/(1+x)^{\frac{1}{2}}$, $(a+bx+cx^2)^{-p}$,

$\{(1+x)^p+(1-x)^q\}/\{(1-x)^p+(1+x)^q\}$, $\{1+x-2\sqrt{(1+x)}\}/\{1-x+2\sqrt{(1+x)}\}$,

$$[1+\{1+(1+x^2)^{\frac{1}{2}}\}^{\frac{1}{2}}]^{\frac{1}{2}}, \qquad \frac{1}{x+}\,\frac{1}{x+}\,\frac{1}{x+}\ldots, \quad x>0.$$

3.01. Prove that the derivative of

$$\log\left\{\frac{\sqrt{(1+x^4)}+x\sqrt{2}}{1-x^2}\right\} \quad \text{is} \quad \frac{(1+x^2)\sqrt{2}}{(1-x^2)\sqrt{(1+x^4)}}.$$

3.1. If $p(x)$ is a polynomial and $p(a) = p'(a) = 0$, show that $p(x)$ has the factor $(x-a)^2$.

3.2. The functions $f(x)$ and $g(x)$ are defined by the following conditions:
$f(x) = -1/2(1+x)$ if $-1 < x \leqslant -\frac{1}{2}$, $f(x) = 2x$ if $-\frac{1}{2} \leqslant x \leqslant \frac{1}{2}$, $f(x) = 1/2(1-x)$ if $\frac{1}{2} \leqslant x < 1$, and $f(x) = 0$ if $x \geqslant 1$ or $x \leqslant -1$.
$g(x) = (2x-1)/2x$ if $x \geqslant 1$, $g(x) = x/2$ if $-1 \leqslant x \leqslant 1$, and $g(x) = -(2x+1)/2x$ if $x \leqslant -1$.
Prove that $f(g(x)) = x$ for all values of x but $g(f(x))$ is different from x outside the interval $[-1, 1]$.

3.3. If the elements of a determinant are differentiable functions, show that the derivative of the determinant is the sum of all the determinants formed by differentiating one row, leaving the other rows unchanged.

3.4. If u_0, u_1, u_2, u_3, u_4 denote a, $ax+b$, $ax^2+2bx+c$, $ax^3+3bx^2+3cx+d$, $ax^4+4bx^3+6cx^2+4dx+e$ prove that

$$u_0^2u_3-3u_0u_1u_2+2u_1^3 \quad \text{and} \quad u_0u_4-4u_1u_3+3u_2^2$$

are independent of x.

3.5. $f(x)$ is differentiable in (a,b), $g(x) = \{f(x)-f(a)\}/(x-a)$ in $[a,b)$ and $g(a) = f'(a)$. Prove that $g(x)$ is continuous in (a,b).

3.6. In the interval (a,b), $|g'(x)| \geqslant \alpha > 0$, $g(\lambda(x))$ is continuous and $a < \lambda(x) < b$; prove that $\lambda(x)$ is continuous in (a,b).

3.7. Show that if $|x| < 1$, $(1+x+x^2)^{-1} = 1-x+x^3-x^4+x^6-x^7+\ldots$ the coefficient of x^n being $1, -1, 0$ according as n leaves the remainder $0, 1, 2$ when divided by 3.

3.8. If $F(x) = x^2$ when $x \leqslant 1$ and $F(x) = 4x-x^2-2$ when $x > 1$, prove that $F(x)$ is differentiable in any interval.

3.81. If $f(x) = x$ when $x \leqslant 1$ and $f(x) = 2-x$ when $x > 1$, prove that $f(x)$ is continuous in any interval, but is not differentiable in any interval which contains the point $x = 1$ in its interior.

3.9. If $F(x)$ is differentiable, with positive non-zero derivative, in (A, B), $F(A) = 0$, and if in the interval $\{F(A), F(B)\}$, $A \leqslant f(x) \leqslant B$, $f(0) = A$ and $f'(x) = 1/F'(f(x))$ prove that $f(x)$, $F(x)$ are inverse functions.

3.91. If $f(x)$ is differentiable and periodic with period a, prove that $f'(x)$ is periodic with period a.

IV

4. Write down the derivatives of the following functions:
e^{x^2}, $\log(1+x^2)$, $2x^2\log x-x^2$, $e^x\log x$, $\log[(1-x+x^2)/(1+x+x^2)]$, $\log\{x+\sqrt{(1+x^2)}\}$, $\log\{x-\sqrt{(x^2-1)}\}$, $xe^x/\log x$, $e^x \operatorname{sh} x$, $\operatorname{ch}(\log x)$, $\log \operatorname{sh} x$, $\log x^{\frac{1}{2}}$, $(\log x)^{\frac{1}{2}}$, $\log_x a$, $a^{\log x}$, x^x, $x^{\log x}$, $(\log x)^x$, $(\log x)^{\log x}$, $e^{\operatorname{sh} x}$, $x^{\operatorname{sh} x}$, $\log(\log x)$, $x\log x\log(\log x)$, $\log_{\log x} x$, $e^{(e^x)}$.

4.1. Find dy/dx when (i) $y = \log_y x$, (ii) $x^y = y^x$, (iii) $e^y = \log_{xy} x$.

4.2. If $f(x)$ is differentiable and $f(x+y) = f(x)f(y)$ for all x and y, prove that either $f(x) = 0$ for all x or $f(x) = e^{kx}$, where k is an arbitrary constant. If $g(xy) = g(x)+g(y)$, $x > 0$, $y > 0$, prove that $g(x) = k\log x$.

4.3. Prove that $\sum n/n! = e$, $\sum n^2/n! = 2e$, $\sum n^3/n! = 5e$, and that $\sum n^k/n!$ is an integral multiple of e.

4.4. If $f(x) = e^{-x}\sum\limits_0^n x^r/r!$ and $g(x) = f(x)+e^{-x}x^{n+1}/(n-m+1)(n!)$, prove that, when $0 < x < m \leqslant n$, $f'(x)$ is negative and $g'(x)$ is positive. Deduce that e^m lies between $\sum\limits_0^n m^r/r!$ and $\left(\sum\limits_0^n m^r/r!\right)+m^{n+1}/(n-m+1)(n!)$, and hence that if $m < 10$ and n is not less than $(p+6)/(1-\log_{10} m)$ then $e^m = \sum\limits_0^n m^r/r!$ correct to p decimal places.

4.5. Prove that $\sum\log(1+1/n)$ is divergent and that $\sum 1/n(\log n)^r$ converges for $r > 1$ and diverges for $r < 1$.

4.51. If $s_n \to 0$, prove that $(1+s_n)^{1/s_n} \to e$ and $(1-s_n)^{1/s_n} \to 1/e$.

4.511. Prove $(1+a/n+b/n^2)^n \to e^a$.

4.52. If $\sum u_r$ is a positive convergent series prove that $\sum\log(1+u_r)$ and $\sum\log(1-u_r)$ are both convergent.

4.521. If $m > 0$ prove that x^m is continuous in $(0, N)$ for any $N > 0$.

4.53. If $a_n > 0$ and $a_n/a_{n+1} < 1+M/n^{1+\lambda}$, $\lambda > 0$, prove that a_n does not tend to zero.

4.6. If $a_n > 0$ and $a_n/a_{n+1} = 1+\beta/n+\theta_n/n^{1+\lambda}$, $|\theta_n| < M$, $\lambda > 0$, then $\sum(-1)^n a_n$ is convergent if $\beta > 0$ and divergent if $\beta \leqslant 0$.

4.7. The Gauss test for convergence. If $a_n > 0$ and

$$a_n/a_{n+1} = 1+\beta/n+\theta_n/n^{1+\lambda}, \qquad |\theta_n| < M, \qquad \lambda > 0,$$

then $\sum a_n$ converges if $\beta > 1$ and diverges if $\beta \leqslant 1$.

4.8. Prove that $(1+x)^m = 1+mx+\binom{m}{2}x^2+\binom{m}{3}x^3+\ldots$ when $x = 1$ and $m > -1$ and when $x = -1$, $m > 0$.

4.81. Show that $x\log(1-1/x)$ is an increasing function.

4.82. If $a_n = (n+1)\log(n+2)-(2n+1)\log(n+1)+n\log n$, prove $\sum a_n$ converges.

4.9. Prove that, for $x > 0$, $(1-x)e^x$ and $(1+x)e^{-x}$ are both decreasing and deduce that, for any x,

$$1-x^2 < e^{-x^2} < 1/(1+x^2).$$

4.91. Prove that

$$\operatorname{sh} x+\operatorname{sh} 2x+\operatorname{sh} 3x+\ldots+\operatorname{sh} nx = \operatorname{sh}\tfrac{1}{2}nx\operatorname{sh}\tfrac{1}{2}(n+1)x/\operatorname{sh}\tfrac{1}{2}x$$

and
$$\tfrac{1}{2}+\operatorname{ch} x+\operatorname{ch} 2x+\ldots+\operatorname{ch} nx = \operatorname{sh}(n+\tfrac{1}{2})x/2\operatorname{sh}\tfrac{1}{2}x.$$

5. Write down the derivatives of the following functions:
$\sin 2x$, $\cos\tfrac{1}{2}x$, $\sin^2 x$, $\sin 2x\cos x$, $\sin^p x\cos px$, $e^x\sin x$, $e^{x^2}\cos 2x$, $\sec x\tan x$, $\sin^{-1}\sqrt{(1-x)}$, $\sin x^2$, $xe^x\sin x$, $x^n e^{px}\cos qx$, $\sin^{-1}\{(3-x)(x-1)\}^{\frac{1}{2}}$, $\cos^{-1}[2\sqrt{\{(a-x)(x-b)\}}/(a-b)]$, $\sin^{-1}[(a+b\cos x)/(a\cos x+b)]$, $\tan^{-1}(\cot x)$,

$\tan^{-1}[(x+\cos\alpha)/\sin\alpha]$, $\tan^{-1}[x\sin\alpha/(1+x\cos\alpha)]$, $\sin^{-1}\{(1+\sin x)/(1+\cos x)\}$, $\log\sin x$, $\log\sec x$, $e^{x\cos\alpha}\sin(x\sin\alpha)$.

5.01. Find the derivatives of the functions $|x^3|$, $|x|$, $|\log(1+x)|$, $|\sin x|$, and $e^{|x|}$, and show that only the first is differentiable in an interval which contains the origin.

5.1. Prove that

$$\cos 3x = 4\cos^3x-3\cos x, \qquad \sin 3x = 3\sin x-4\sin^3x,$$
$$\tan 3x = (3\tan x-\tan^3x)/(1-3\tan^2x).$$

5.11. Prove that

$$\sin\tfrac{1}{4}\pi = \cos\tfrac{1}{4}\pi = \frac{1}{\sqrt{2}}, \qquad \sin\tfrac{1}{6}\pi = \cos\tfrac{1}{3}\pi = \tfrac{1}{2},$$
$$\cos\tfrac{1}{12}\pi = \frac{1+\sqrt{3}}{2\sqrt{2}}, \qquad \sin\tfrac{1}{12}\pi = \frac{\sqrt{3}-1}{2\sqrt{2}}.$$

5.12. Prove that $\frac{1}{2}+\cos x+\cos 2x+...+\cos nx = \frac{1}{2}\sin(n+\frac{1}{2})x/\sin\frac{1}{2}x$, and $\sin x+\sin 2x+...+\sin nx = \sin\frac{1}{2}nx\sin\frac{1}{2}(n+1)x/\sin\frac{1}{2}x$.

5.2. Prove the formulae

$$\sin nx = \binom{n}{1}\cos^{n-1}x\sin x-\binom{n}{3}\cos^{n-3}x\sin^3x+\binom{n}{5}\cos^{n-5}\sin^5x-...,$$

$$\cos nx = \cos^nx-\binom{n}{2}\cos^{n-2}x\sin^2x+\binom{n}{4}\cos^{n-4}x\sin^4x-...$$

and deduce that $\cos nx$ and $\sin(n+1)x/\sin x$ are polynomials in $\cos x$ of the nth degree.

5.201. Polynomials $C_n(x)$, $S_n(x)$ are defined for all values of x and all positive integral values of n by the equations

$$C_0(x) = 1, \qquad C_1(x) = x, \qquad C_{n+2}(x) = 2xC_{n+1}(x)-C_n(x)$$
and $$S_0(x) = 1, \qquad S_1(x) = 2x, \qquad S_{n+2}(x) = 2xS_{n+1}(x)-S_n(x).$$

Prove that $C_n(\cos\theta) = \cos n\theta$ and $S_n(\cos\theta) = \sin(n+1)\theta/\sin\theta$.

5.202. Prove that $\sum_1^n x^nC_n(x) = x^{n+1}S_{n-1}(x)$ and deduce the sum to n terms of the series $\sum\cos^n\theta\cos n\theta$.

5.21. Show that

$$\cos nx-\cos n\alpha = 2^{n-1}(\cos x-\cos\alpha)\left\{\cos x-\cos\left(\alpha+\frac{2\pi}{n}\right)\right\}\times$$
$$\times\left\{\cos x-\cos\left(\alpha+\frac{4\pi}{n}\right)\right\}...\left\{\cos x-\cos\left(\alpha+\frac{2(n-1)\pi}{n}\right)\right\},$$

and also that

$$C_n(u)-C_n(\cos\alpha) = 2^{n-1}(u-\cos\alpha)(u-\cos[\alpha+2\pi/n])\times$$
$$\times(u-\cos[\alpha+4\pi/n])...(u-\cos[\alpha+2(n-1)\pi/n]).$$

Prove

$$x^{2n}-2x^n\cos n\alpha+1 = \{x^2-2x\cos\alpha+1\}\left\{x^2-2x\cos\left(\alpha+\frac{2\pi}{n}\right)+1\right\}\times$$
$$\times\left\{x^2-2x\cos\left(\alpha+\frac{4\pi}{n}\right)+1\right\}...\left\{x^2-2x\cos\left(\alpha+\frac{2(n-1)\pi}{n}\right)+1\right\}.$$

Deduce the factors of x^n-1 and x^n+1.

5.22. Prove that $$D_x(x^n C_n(x)) = nx^{n-1}S_n(x).$$

5.3. Prove that for any value of x, $\sin^4 x$ lies between $x^4-32x^6/45$ and $x^4+2x^6/45$.

5.301. Prove that $\cos^n \dfrac{x}{n} \to 1$.

5.31. If $0 < |x| \leqslant \pi$, prove $\{(\sin x)/x\}^3 > \cos x$.

5.32. Prove $\{1/\sin(1/n)\}-n \to 0$.

5.33. Prove that, for $0 < x \leqslant \frac{1}{2}\pi$,

$$\text{(i)}\quad x\cos x < \sin x, \qquad \text{(ii)}\quad \cos x < e^{-\frac{1}{2}x^2}.$$

5.34. Prove that the series $\sum(-1)^n n\{1-\cos(x/n)\}$ converges for all values of x.

5.4. If $f(x)$ is differentiable and $f(x)+f(y) = f((x+y)/(1-xy))$ prove that $f(x) = 0$ or $f(x)$ is a constant multiple of $\tan^{-1}x$.

5.5. Show that the derivative of each of the four functions

$$\log(\sec x+\tan x), \qquad \mathrm{ch}^{-1}\sec x, \qquad \mathrm{sh}^{-1}\tan x, \quad \text{and} \quad \mathrm{th}^{-1}\sin x$$

is $\sec x$, and hence prove that they are all equal.

Show further that $\operatorname{sech} x$ is the derivative of each of $\cos^{-1}\operatorname{sech} x$, $\sin^{-1}\mathrm{th}\, x$, $\tan^{-1}\mathrm{sh}\, x$ and hence that these, too, are equal.

5.6. Prove that if $0 \leqslant r < 1$,

$$1+r\cos\theta+r^2\cos 2\theta+r^3\cos 3\theta+\ldots = \frac{1-r\cos\theta}{1-2r\cos\theta+r^2},$$

$$r\sin\theta+r^2\sin 2\theta+r^3\sin 3\theta+\ldots = \frac{r\sin\theta}{1-2r\cos\theta+r^2}.$$

5.7. Prove that $\displaystyle\sum_{n\geqslant 1}\frac{\cos nx}{n^2}$ is continuous in any interval.

VI

6. Resolve the following expressions into their partial fractions:

$x/(x^2-1)$, $x^3/(x+1)(x+2)(x+3)$, $(x^2-x+2)/(x^4-5x^2+4)$, $(x+1)/(x-1)^2(x+2)$, $x/(x-1)^2(x^2+x+1)$, $(x-1)/(x-3)(x^2+4)$, $(x-1)/(x-3)(x^2+4)^2$, $(x^2+2x+4)/(x+1)(x^2+1)^2$, $(x^4+x^2+1)/x^2(x^2+1)^2$, $(x+9)/(x-1)^2(x^2+4x+5)$, $(x+2)/(x-1)^2(x^2-2x+2)$, $4/(1-x^2)^2$, $(x+1)/(x^2+1)(x^2+4)(x^2+9)$, $(x+1)^2/(x^4+x^2+1)$, $(x^2+2x+3)/(x+1)(x^2+1)(x^2+2)$, $(x+1)/(x^4+1)$, $(x^2+1)^2/(x^8+x^4+1)$, $(1+x^2)/(1-2x^2\cos 2\theta+x^4)$.

VII

7. If $y = x^{\frac{1}{3}}+x^{\frac{2}{3}}$ prove that $9x^2\dfrac{d^2y}{dx^2}+2y = 0$ and hence that

$$x^2\frac{d^{n+2}y}{dx^{n+2}}+2nx\frac{d^{n+1}y}{dx^{n+1}}+(n-\tfrac{1}{3})(n-\tfrac{2}{3})\frac{d^ny}{dx^n} = 0.$$

7.01. If $2y = x\left(1+\dfrac{dy}{dx}\right)$ prove that d^2y/dx^2 is constant.

7.02. If $f(x)$, $g(x)$ are differentiable three times, with first derivatives $f'(x)$, $g'(x)$ respectively, and if $f'(x)g'(x) = 1$ and $h(x) = f(x)g(x)$, prove that

$$\frac{h^3(x)}{h(x)} = \frac{f^3(x)}{f(x)} + \frac{g^3(x)}{g(x)}.$$

7.1. Prove that if $y = 2/(1+x^2)$ then $\dfrac{d^3y}{dx^3} = 6xy^3(1-x^2y)$.

7.11. If $ky = \sin(x+y)$, where k is constant, prove that

$$\frac{d^2y}{dx^2} = -y\left(1+\frac{dy}{dx}\right)^3.$$

Assuming that, when $k > 1$, y can be expanded in a series of ascending powers of x, calculate the coefficients up to that of x^4. What is the part played by the restriction on k?

7.2. Find the nth derivatives of

(i) $(x+1)/(x-1)^2(x+2)$, (ii) x^4e^x, (iii) $\sin^3 x \cos x$.

7.21. If $y = \frac{1}{2}x^2$ and $z = \sin y$, prove that d^nz/dx^n is of the form $x(P_n \sin y + Q_n \cos y)$ if n is odd, and of the form $P_n \sin y + Q_n \cos y$ if n is even, where P_n and Q_n are polynomials in y. Show further that

$$P_{2n+2} = \frac{dP_{2n}}{dy} - Q_{2n} + 2y\left\{\frac{d^2P_{2n}}{dy^2} - 2\frac{dQ_{2n}}{dy} - P_{2n}\right\}.$$

7.211. Prove that the nth derivative of $\tan x$ is a polynomial in $\tan x$ of the $(n+1)$th degree.

7.22. If $y = e^{x\cos\alpha}\sin(x\sin\alpha)$ and y_n denotes d^ny/dx^n prove that

$$y_{n+2} - 2y_{n+1}\cos\alpha + y_n = 0.$$

Assuming that y can be expanded in a convergent series, prove that

$$y = \sum \frac{x^n}{n!}\sin n\alpha.$$

7.23. If $p(x)$ is a polynomial and $p(a) = p'(a) = \dots = p^n(a) = 0$, prove that $p(x)$ is divisible by $(x-a)^{n+1}$.

7.3. If $f'(a) = f'(b) = 0$ and $f'(x) \neq 0$ in $[a, b]$ then $f(a)$, $f(b)$ cannot both be maxima nor both minima.

7.31. If $f(x)$ is differentiable in (a, b) and if $f'(x)$ vanishes at only a finite number of points in (a, b), then between any two points in $[a, b]$ where $f(x)$ is maximum there is a point where $f(x)$ is minimum, and between two minimum values, a maximum.

7.4. Find the maximum and minimum values of the functions

(i) $x^2/(x-a)(x-b)$, $a > b > 0$; (ii) $e^{a\sin^2x + b\cos^2x}$, $a > b > 0$;
(iii) $(x-a)(x-b)/(x-c)$, $a < b < c$; (iv) $(x^2+5x-5)/(x^2-3x+2)$;
(v) $x(2-9\log_x e)$; (vi) $(\log x)^3 - \log x^3$; (vii) $(x^6-36x^3-16)/x^2$.

7.41. Show that if $a > b > 0$ the minimum value of $(a-b)x/(x+a)(x+b)$ exceeds the maximum by $4\sqrt{(ab)}/(a-b)$.

7.411. Prove that the function $x^2(x^2-4)$ has two minimum values and one maximum value.

7.42. Determine all the maximum and minimum values of the function $\sin^m x/\sin m(x-\alpha)$, where m is a positive integer and $0 < \alpha < \pi/m$.

7.43. If $l > m > 0$, find the maximum and minimum values of the function

$$\sin^2 x \cos^2 x\{l(\sin x - \cos x)^2 + m(\sin x + \cos x)^2\}.$$

7.44. Find the maximum and minimum values of the function

$$\sin^3 x \cos 3x.$$

7.5. A triangle ABC has a right angle at C, and the product of the sides AB, BC is constant. Prove that $AC+3BC$ is a minimum when $AC = 2BC$ and a maximum when $AC = BC$.

7.6. $t_r, u_r, v_r, w_r, \ldots$ denote the rth derivatives of the $n+1$ functions of x, $t_0, u_0, v_0, w_0, \ldots$ respectively. Both the determinants of the $(n+1)$th order $\Delta = (t_0 u_1 v_2 w_3 \ldots)$ and $\Gamma = (t_1 u_2 v_3 w_4 \ldots)$ are zero for all values of x but the minor of t_0 in Δ is not zero. Prove that the minor in Δ of each of $u_0, v_0, w_0, \ldots$ is a constant multiple of the minor of t_0.

7.7. If $y = \tan^{-1}x$ and y_n is the nth derivative of y with respect to x, prove that $(1+x^2)y_{n+1} + 2nxy_n + n(n-1)y_{n-1} = 0$, and assuming that for $-1 < x \leqslant 1$, y can be expanded in a convergent series of ascending powers of x, prove that

$$\tan^{-1}x = x - \tfrac{1}{3}x^3 + \tfrac{1}{5}x^5 - \ldots, \quad -1 < x \leqslant 1.$$

7.71. If $y = \sin^{-1}x$ and y_n is the nth derivative of y, prove that

$$(1-x^2)y_{n+2} - (2n+1)xy_{n+1} - n^2 y_n = 0.$$

Assuming that, for $|x| < 1$, y can be expanded in a convergent series of ascending powers of x, prove that

$$\sin^{-1}x = \sum \frac{1.3.5\ldots(2n-1)}{2.4.6\ldots 2n}\,\frac{x^{2n+1}}{2n+1}, \quad |{:} \qquad 1.$$

7.8. Deduce the expansions of Examples 7.7, 7.71 from the binomial theorem.

7.9. Prove that if $ax^2+2bx+c$ has no linear factors then λ is a maximum or minimum value of the function $\{(px+q)^2/(ax^2+2bx+c)\}+\lambda$ according as $aq^2-2bpq+cp^2$ is negative or positive.

7.91. Show that the values of λ for which $Ax^2+2Bx+C-\lambda(ax^2+2bx+c)$ is a perfect square are the maximum and minimum values of

$$(Ax^2+2Bx+C)/(ax^2+2bx+c),$$

provided $ax^2+2bx+c$ has no linear factors.

7.92. If $P(x)$ is a polynomial of the nth degree, prove that

$$P(a+h) = P(a) + hP'(a) + \frac{h^2}{2!}P''(a) + \ldots + \frac{h^n}{n!}P^n(a).$$

7.921. If $f(x) = \sum a_n x^n$, $g(x) = \sum b_n x^n$, the series being convergent for $|x| < R$, and if there is an r such that $f(x) = g(x)$ for $|x| < r < R$, prove that $f(x) = g(x)$ for $|x| < R$.

VIII

8. Find the indefinite integrals of the following functions:

$x^{\frac{2}{3}}$, $\sqrt[3]{x^2}$, $\sqrt{x}+1/\sqrt{x}$, $\sin\frac{1}{2}x$, $\sin x\cos x$, x^2e^{3x}, $e^{ax}\sin bx$, $1/x(1+x^2)$, $e^x\cos^2 x$, $\sec^3 x$, $x\sqrt{(2-x^2)}$, $(x-2)/\sqrt{(x^2-4x+5)}$, $x\sin x$, $x\sin^{-1}x$, $\tan^3 x$, $\sin^{-1}x$, $(x^3-x+2)/(x^4-5x^2+4)$, $x/(a^4+x^4)$, $x^2/(x-1)^5$, $1/(x^4+1)$, $x^2/(x^4+1)$, $x\sec^2 x$, $e^{2x}\cos^2 3x$, $1/(5-4\cos x)$, $(x-1)/(x-3)(x^2+4)^2$,

$\cos x/(1-\cos\alpha\cos x)$, $\sin^2 x/(1-\cos\alpha\cos x)^2$, $x\,\mathrm{sh}^{-1}x/(1+x^2)^{\frac{3}{2}}$, $\log(1+x)/(1-x)$, $(2x^2+3x+4)/(4x^2+9)^2$, $x\{(2-x)/(3+x)\}^{\frac{1}{2}}$, $x^2\{(1-x)/(1+x)\}^{\frac{1}{2}}$, $x(x+2)^{\frac{1}{2}}$, $1/(x-1)\sqrt{(x^2+1)}$, $(1+x)/(x^2+1)^2$, $x^2\sin x$, $x^2(\log x)^2$, $xe^{ax}\,\mathrm{sh}\,ax$, $xe^{ax}\sin ax$, $\sqrt{(x^2+1/x^2)}$ (take $x^4 = u^2-1$), $(x+9)/(x-1)^2(x^2+4x+5)$, $(6-2x)/(x^2-2x+5)^2$, $1/(x-1)\sqrt{(x^2-2x+5)}$, $(x+1)/(x^2+1)(x^2+4)(x^2+9)$, $(x^2+x+1)/x(x^2+1)$, $x/(x+1)(x^2+1)^2$, $1/x\sqrt{(x^2-4)}$, $1/(x^2+1)^3$, $1/(3+5\cos x)$, $\sqrt{2}/(\sin x+\cos x)$, $(x+1)/(x^2+4)\sqrt{(x^2+9)}$, $(1+x)/(4x^2+1)\sqrt{(9x^2-4)}$, $(x-1)/(x^2+1)\sqrt{(x^2-1)}$, $1/(5x^2+12x+8)\sqrt{(5x^2+2x-7)}$, $(x-1)/(2x^2-6x+5)\sqrt{(7x^2-22x+19)}$.

Evaluate the integrals:

8.01. $\int [1/\{\sqrt{(x+2)}+\sqrt{(x+1)}\}]\,dx$; 8.02. $\int dx/x(1+x^5)$;

8.03. $\int \{(3\sin x+4\cos x)/(\sin x+\cos x)\}\,dx$; 8.04. $\int dx/\sqrt{(2e^x-1)}$;

8.05. $\int [\sqrt{(20x-15)}/\{35-8\sqrt{(16-5x)}\}]\,dx$; 8.06. $\int x^5(1-x)^{-3}\,dx$.

8.07. If $f'(x)$, $g'(x)$ are the derivatives of $f(x)$, $g(x)$ prove that

$$\frac{1}{2}\int \frac{1}{\sqrt{(fg)}}\frac{fg'-f'g}{f-g}\,dx = \log\frac{\sqrt{f}+\sqrt{g}}{\sqrt{(f-g)}}.$$

8.08. If $u_n = \int (1-x^4)^{-n}\,dx$ prove that

$$4nu_{n+1} = (4n-1)u_n+x(1-x^4)^{-n}, \quad n \geqslant 1,$$

and hence evaluate the integrals

$$\int x^4(1-x^4)^{-3}\,dx, \qquad \int x^8(1-x^4)^{-4}\,dx.$$

8.1. Integrate the function $(1+x^2)^{\frac{1}{4}}x^{-\frac{3}{2}}$ by means of the substitution $1/x^2 = t^4-1$.

8.2. Evaluate $\int (\log x)^2\,dx$ and $\int (\log x)^3\,dx$ and find the relation between $\int (\log x)^{k+1}\,dx$ and $\int (\log x)^k\,dx$.

8.21. If $g^*(t)$ is the inverse function of $g(t)$ show that under the transformation $t = g^*(x)$ the integral $\int f(g(t))g'(t)\,dt$ becomes $\int f(x)\,dx$.

8.3. Show that if $y^2 = (x^2-8x+10)/(3x^2-10x+9)$ then

$$\sqrt{14}\int (x-4)\,dx/(3x^2-10x+9)\sqrt{(x^2-8x+10)}$$
$$= 3\int dy/\sqrt{(\tfrac{1}{2}-y^2)}-2\sqrt{2}\int dy/\sqrt{(y^2+2)}$$

and hence evaluate the integral (use Example 7.91).

8.31. Evaluate $\int dx/(5x^2+12x+8)\sqrt{(5x^2+2x-7)}$ by the substitution

$$y^2 = (5x^2+2x-7)/(5x^2+12x+8).$$

8.4. Evaluate the integrals

(i) $\displaystyle\int \frac{x+\sqrt{(x^2-a^2)}}{x-\sqrt{(x^2-a^2)}}\frac{dx}{\sqrt{(x^2-a^2)}}$,

(ii) $\displaystyle\int \left(1+\frac{x}{\sqrt{(x^2-a^2)}}\right)^{-2}dx$

by the substitution $x+\sqrt{(x^2-a^2)} = t$.

IX

9. Evaluate the integrals:

$\int_0^{1/\sqrt{2}}\{1/\sqrt{(1-x^2)}\}\,dx$, $\int_1^2\{x/(1+x^2)^2\}\,dx$, $\int_2^4\{(\log x)/x\}\,dx$, $\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} dx/(5-4\cos x)$,

$\int_0^1 dx/(\sqrt{2}+\cos x-\sin x)$, $\int_3^5 dx/(x-2)\sqrt{(x^2-4x+3)}$, $\int_{\frac{3}{2}}^{\frac{11}{2}} dx/(x-1)\sqrt{(x^2-2x+5)}$,

$\int^{3}\{(x+2)/(x-1)^2(x^2-2x+2)\}\,dx$, $\int_0^3(6-2x)\,dx/(x^2-2x+5)^2$, $\int_{\frac{3}{4}}^{\frac{4}{3}} dx/x\sqrt{(1+x^2)}$,

$\int_0^{\frac{1}{4}\pi}\sqrt{2}\,dx/(\sin x+\cos x)$, $\int_{-\frac{1}{2}}^{\frac{1}{2}}\{(x+1)^2/(x^4+x^2+1)\}\,dx$, $\int_{\frac{1}{3}\pi}^{\frac{1}{2}\pi}\operatorname{cosec}^5\tfrac{1}{2}x\,dx$,

$\int_0^1\{x/(1-x)\}^{\frac{1}{2}}\,dx$, $\int_1^2 dx/x(1+x^5)$, $\int_0^a dx/\{x+\sqrt{(a^2-x^2)}\}$, $\int_3^4\{(x-3)/(x-2)\}^{\frac{1}{2}}\,dx$,

$\int_0^{\sqrt{2}} x\sqrt{(2-x^2)}\,dx$, $\int_0^{\frac{1}{2}\pi}\sin^8 x\,dx$, $\int_2^3 x\{(4-x)/(2x-3)\}^{\frac{1}{2}}\,dx$, $\int_0^1\sqrt{\{(3-x)(3-2x)\}}\,dx$,

$\int_1^2 dx/\sqrt{(2e^x-1)}$, $\int_{\frac{3}{2}}^2\{(5-x)/(x-1)\}^{\frac{1}{2}}\,dx$, $\int_0^2\{(3-x)/(2-x)\}^{\frac{1}{2}}\,dx$, $\int_1^2 x^7\,dx/(x^4+2)$,

$\int_2^3\{\log x/(x+1)^2\}\,dx$, $\int_0^{\frac{1}{2}\pi} dx/(1+\sin x)^2$, $\int_3^4(4x+3)\,dx/\sqrt{(x^2+10x+26)}$,

$\int_1^2 dx/\{\sqrt{(2x-1)}+\sqrt{(5-2x)}\}$, $\int_1^2 x\,\mathrm{sh}^{-1}x\,dx/(1+x^2)^{\frac{3}{2}}$, $\int_0^{\sin^{-1}\frac{4}{5}} dx/(5\cos x-3)^{\frac{1}{2}}\cos\tfrac{1}{2}x$,

$\int_0^1(1-x^{\frac{1}{2}})^{\frac{1}{2}}x^{-\frac{1}{2}}\,dx$, $\int_{\frac{1}{2}}^4\{1/x\sqrt{(x^2+8x+1)}\}\,dx$.

9.01. Evaluate the integrals $\int_{\frac{1}{4}}^{\frac{1}{2}}\{1/(1+x^2)\}\,dx$ and $\int_{\frac{1}{4}}^{\frac{1}{2}}\{x^2/\sqrt{(1-x^2)}\}\,dx$ to three places of decimals.

9.02. Prove that $\displaystyle\int_0^1\frac{1}{\sqrt{\{x(1-x)\}}}\,dx=\pi.$

9.1. Prove that

$$\int_0^{\frac{1}{2}\pi}\sin^{n-1}\theta\sin(n+1)\theta\,d\theta=(1/n)\sin(\tfrac{1}{2}n\pi)$$

and
$$\int_{\pi/n}^{\frac{1}{2}\pi}\{\sin(n-1)\theta/\sin^{n+1}\theta\}\,d\theta=-(1/n)\sin(\tfrac{1}{2}n\pi).$$

9.101. By means of the transformation $x=\sin t$, prove that

$$\int_{\frac{1}{2}}^{\sqrt{\frac{3}{4}}}2x\,dx=\int_{-7\pi/6}^{7\pi/3}\sin 2t\,dt.$$

9.11. Show that, if $n>1$,

$$\int_0^\infty x^n e^{-x^2}\,dx=\tfrac{1}{2}(n-1)\int_0^\infty x^{n-2}e^{-x^2}\,dx,$$

and hence that if n is a positive integer $2\int_0^\infty x^{2n+1}e^{-x^2}\,dx=n!$.

9.12. Prove that

$$\int_0^\infty dx/(1+e^x)^p(1+e^{-x}) = 2^{-p}/p, \quad p > 0, \qquad \int_0^1 (1-x^2)^n\, dx = 2n!/(2n+1)!,$$

$$\int_0^{2a} x\, dx/\sqrt{(2ax-x^2)} = \pi a, \qquad \int_0^\infty (1+x^2)^{-n}\, dx = \{(2n-3)!/(2n-2)!\}(\tfrac{1}{2}\pi),$$

$$\int_0^{\frac{1}{2}\pi} \{(3\sin x+4\cos x)/(\sin x+\cos x)\}\, dx = 7\pi/4.$$

9.13. Prove that

$$\int_0^{\frac{1}{2}\pi} \{\sin(2n+1)x/\sin x\}\, dx = \tfrac{1}{2}\pi$$

and

$$\int_0^{\frac{1}{2}\pi} (\sin nx/\sin x)^2\, dx = \tfrac{1}{2}n\pi.$$

9.2. Show that

$$\int_{-1}^1 \sin\alpha\, dx/(1-2x\cos\alpha+x^2) = \tfrac{1}{2}\pi \quad \text{if} \quad 2n\pi < \alpha < (2n+1)\pi$$
$$= -\tfrac{1}{2}\pi \quad \text{if} \quad (2n-1)\pi < \alpha < 2n\pi.$$

9.201. Prove that

$$\int^1 \frac{}{(x+1)\sqrt{\{(1-x)(1-x\cos 4\alpha)\}}}\, dx = \sec 2\alpha \log\cot\alpha, \quad 0 < \alpha < \tfrac{1}{4}\pi.$$

9.21. Evaluate the integral

$$\int_{[illegible]}^{[illegible]} dx/x^3\sqrt{(1-x^2)}.$$

9.3. If $f(x)$ is symmetrical about the point $x = l$ prove that

$$\int_0^{2l} xf(x)\, dx = 2l\int_0^l f(x)\, dx$$

and hence that

$$\int_0^\pi x\log_e \sin x\, dx = -\tfrac{1}{2}\pi^2 \log_e 2.$$

9.31. Prove that

$$\int_0^a \sec x \sec(a-x)\, dx = 2\operatorname{cosec} a \log\sec a$$

and deduce the value of

$$\int_0^a x\sec x\sec(a-x)\, dx.$$

9.32. Prove that

$$\int_2^3 dx/\{\sqrt{(5-x)}+\sqrt{(x-1)}\} = \sqrt{2}\log\tan(5\pi/24)+\sqrt{3}-1.$$

9.33. Show that the generalized integral $\int_0^\pi \frac{\sin x}{x}\,dx$ exists.

9.34. Prove that

$$\int_0^1 \log(1+x)\,dx/(1+x^2) = \tfrac{1}{8}\pi\log 2.$$

9.35. If $\phi(x)$ is positive and steadily decreasing and $\int_a^\infty \phi(x)\,dx$ exists, prove that $n\phi(n) \to 0$.

9.36. If $\int_0^1 f(xt)\,dt = 0$ for all values of x, prove $f(x) = 0$

9.4. Show that $\int_0^1 x^{n-1}\,dx/(1+\surd x)$ lies between $1/n(2n+1)$ and $1/n$ and hence by evaluating the integral prove that $\log 2 = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}$.

9.401. If $B > A > 0$ show that $\left|\int_A^B (\sin x/x)\,dx\right| < 2/A$ and deduce that $\int_a^\infty (\sin x/x)\,dx$ exists, $a > 0$.

9.41. If $l(x) = \int_1^x t^{-1}\,dt$, $x > 0$, prove that $l(x^{-1}) = -l(x)$, $l(xy) = l(x)+l(y)$, $l(x^y) = yl(x)$ without assuming the properties of the logarithmic or exponential function.

9.5. Prove that

$$\left(1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\dots+\frac{x^k}{k!}\right)e^{-x} = \sum_{r=0}(-1)^{r+1}x^{k+r+1}/(k!)(r!)(k+r+1).$$

9.51. Show that the series $\operatorname{sh}\alpha/e^\beta-\operatorname{sh}3\alpha/3e^{3\beta}+\operatorname{sh}5\alpha/5e^{5\beta}-\dots$ is convergent if $-\beta < \alpha < \beta$, and find its sum.

9.52. If α and β are the roots of the equation $x^2-ax+b = 0$, prove that

$$\log(1+ax+bx^2) = (\alpha+\beta)x-\tfrac{1}{2}(\alpha^2+\beta^2)x^2+\tfrac{1}{3}(\alpha^3+\beta^3)x^3-\dots$$

and state the range of values of x for which the expansion is valid.

Hence sum the series

$$x\operatorname{ch}\theta-\tfrac{1}{2}x^2\operatorname{ch}2\theta+\tfrac{1}{3}x^3\operatorname{ch}3\theta-\dots$$

and say for what values of x the sum exists.

9.53. Show that, for $|x| < 1$,

$$x\cos\theta-\tfrac{1}{2}x^2\cos2\theta+\tfrac{1}{3}x^3\cos3\theta-\dots = \tfrac{1}{2}\log(1+2x\cos\theta+x^2)$$

and

$$x\sin\theta+\tfrac{1}{3}x^3\sin3\theta+\tfrac{1}{5}x^5\sin5\theta+\dots = \tfrac{1}{2}\tan^{-1}\{2x\sin\theta/(1-x^2)\}.$$

9.531. Prove that, for $|r| < 1$,

$$r\cos\theta+\tfrac{1}{2}r^2\cos2\theta+\tfrac{1}{3}r^3\cos3\theta+\dots = -\tfrac{1}{2}\log(1-2r\cos\theta+r^2),$$

$$r\sin\theta+\tfrac{1}{2}r^2\sin2\theta+\tfrac{1}{3}r^3\sin3\theta+\dots = \tan^{-1}\{r\sin\theta/(1-r\cos\theta)\}.$$

9.532. Prove that, for all values of r and θ,

$$\mathbf{1+r\cos\theta+\frac{r^2}{2!}\cos 2\theta+\frac{r^3}{3!}\cos 3\theta+\ldots = e^{r\cos\theta}\cos(r\sin\theta),}$$

$$\mathbf{r\sin\theta+\frac{r^2}{2!}\sin 2\theta+\frac{r^3}{3!}\sin 3\theta+\ldots = e^{r\cos\theta}\sin(r\sin\theta).}$$

9.54. If $\phi(x)$ is positive, continuous, and steadily decreasing as x increases, show that $\sum\phi(n)$ converges and diverges together with the sequence $\int\limits_a^n \phi(x)\,dx$. Prove that $\sum n^k e^{-n}$ converges.

9.6. Prove that

$$\sin x+\tfrac{1}{2}\sin 2x+\tfrac{1}{3}\sin 3x+\ldots = \tfrac{1}{2}(\pi-x), \quad 0 < x < 2\pi.$$

9.61. Prove that

$$\sum_{n\geq 1}\frac{\cos nx}{n^2} = \frac{x^2}{4}-\frac{\pi x}{2}+\frac{\pi^2}{6}, \quad 0 \leq x \leq 2\pi.$$

9.62. If $\phi(x)$ is a polynomial of the nth degree, prove that $\int\limits_0^c \frac{\phi(x)}{\sqrt{(cx-x^2)}}\,dx$ is also a polynomial of the nth degree.

9.63. Find the derivative of the function $\int\limits_a^x \phi(y)\,dy \int\limits_x^b \psi(y)\,dy$ and hence show that

$$\int\limits_a^b \psi(x)\left\{\int\limits_a^x \phi(y)\,dy\right\}dx = \int\limits_b^a \phi(x)\left\{\int\limits_b^x \psi(y)\,dy\right\}dx.$$

9.64. If $f(x)$ is periodic with period a, and if $\phi(x) = f(x)-\frac{1}{a}\int\limits^a f(x)\,dx$, prove that $\int\limits_0^x \phi(t)\,dt$ is periodic, with period a.

9.7. Show that

$$\int\limits_0^{\frac{1}{2}\pi}\sin^{2n+1}x\,dx < \int\limits_0^{\frac{1}{2}\pi}\sin^{2n}x\,dx < \int\limits_0^{\frac{1}{2}\pi}\sin^{2n-1}x\,dx$$

and deduce Wallis's formula

$$\tfrac{1}{2}\pi = \frac{2.2}{1.3}\cdot\frac{4.4}{3.5}\cdot\frac{6.6}{5.7}\cdots\frac{2n.2n}{(2n-1)(2n+1)}\cdots.$$

9.8. Prove that the generalized integral $\int\limits_0^\infty e^{-x^2}\,dx$ exists and that its value is $\sqrt{\pi}/2$

9.9. Evaluate the integral $\int\limits_{-1}^{2} x\,dx/x^{\frac{1}{2}}$ by the substitution $x^2 = y$ and the integral $\int\limits_1^7 (x^2-6x+13)\,dx$ by the substitution $y = x^2-6x+13$.

9.91. Prove that

$$n!/\sqrt{(2\pi n)}(n/e)^n \to 1 \quad \text{(Stirling's formula).}$$

9.92. Prove
$$\int_0^1 \{1/\sqrt{(\log(1/x))}\}\,dx = \sqrt{\pi}.$$

9.93. Show that
$$e^{-n}\left\{1+\frac{n}{1!}+\frac{n^2}{2!}+\frac{n^3}{3!}+\ldots+\frac{n^n}{n!}\right\} \to \tfrac{1}{2}.$$

9.94. If $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}\,dt$ prove that $\Gamma(a)$ is defined for $a > 0$. Show further that, if $a > 0$, $\Gamma(a+1) = a\Gamma(a)$ and $\Gamma(\tfrac{1}{2}) = \sqrt{\pi}$.

Deduce that
$$\int_0^{\frac{1}{2}\pi} \sin^m x \cos^n x\,dx = \tfrac{1}{2}\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)\Big/\Gamma\left(\frac{m+n+2}{2}\right),$$
m and n being positive integers.

X

10. Find the lengths of the subtangent and subnormal of the cycloid $x = a(\theta-\sin\theta)$, $y = a(1-\cos\theta)$ and hence show that the normal at the point θ passes through the point $x = a\theta$ on the x-axis.

10.1. Find the area of a loop of each of the following curves:

10.11. $x = 2a\sin t$, $y = 2a\sin 2t$;

10.12. $x : y : 1 = (1-t^2) : t(1-t^2) : (1+t^2)$,

10.13. $x = at/(1+t^3)$, $y = at^2/(1+t^3)$;

10.14. $r = 5a\sin^2\theta\cos^2\theta/(\sin^5\theta+\cos^5\theta)$;

10.15. $y^3-3xy+2x^3 = 0$; 10.16. $r = 2\cos\theta-1$.

10.2. $f(x)$ is a polynomial and $y = mx+c$ is the tangent to $y = f(x)$ at $x = x_0$; show that the polynomial $f(x)-mx-c$ is divisible by $(x-x_0)^2$.

10.21. If $f(x) = ax+6b^2/x^2$, $b \neq 0$, prove that the curve $y = e^{-f(x)}$ has at least two points of inflexion.

10.3. Find the perimeter of, and the area bounded by, each of the closed curves:

10.31. $x = \cos^3 t$, $y = \sin^3 t$;

10.311. $x = te^{-2t}$, $y = t^2e^{-2t}$, $t \geqslant 0$;

10.32. $x = \cos t(1+\cos t)$, $y = \sin t(1+\cos t)$.

10.33. Find the length of the curve $x = 2\,\mathrm{ch}\,t$, $y = \mathrm{sh}\,t\,\mathrm{ch}\,t-1$ between the two points at which $x = 8$, and of the curve $y = 2\tan^2 x+\tfrac{1}{2}\log\cos^2 x$ between the points $x = 0$ and $x = \tfrac{1}{4}\pi$.

10.34. By writing $y = a\sin^3 t$ obtain a parametric representation of the curve
$$4(x^2+y^2)-a^2 = 3a^{\frac{4}{3}}y^{\frac{2}{3}},$$
and hence prove the area contained by the curve is $3\pi a^2/4$, and find the length of the curve.

10.35. Find the area between the curves $y = (x^2-1)^2$ and $y = 9x^2/4$.

10.36. Show that the length of a quadrant of the curve $(x/a)^{\frac{2}{3}}+(y/b)^{\frac{2}{3}} = 1$ is $(a^2+ab+b^2)/(a+b)$.

10.37. In the cycloid $x = a(t+\sin t)$, $y = a(1-\cos t)$, prove that $s = 4a\sin\psi$, where s is the arc length, measured from the origin, and ψ is the inclination of the tangent.

10.4. If s is the arc and A the sector area of the curve

$$(a+r)/(a-r) = e^{\theta},$$

both measured from $\theta = 0$, prove that

$$s+r = a\theta, \qquad A+ar = \tfrac{1}{2}a^2\theta.$$

10.41. In the semi-cubical parabola $y = ax^{\frac{3}{2}}$ find the relation between the arc s measured from the origin to a variable point P and the angle ψ between the tangent at O and the tangent at P.

Prove also that if the normal to the parabola at P cuts the axes in G, H, the radius of curvature at P is proportional to HP^4/PG.

10.42. Prove that, if the tangent to a curve makes an angle ψ with a fixed line and is at a distance p from a fixed point, the radius of curvature is

$$|p+d^2p/d\psi^2|.$$

Prove that in a curve for which

$$p = \sin\psi \log(\sec\psi+\tan\psi)-1$$

the arc s, measured from a suitable point, is equal to $\tan\psi$, and the reflection of the centre of curvature at any point in the tangent at that point lies on a fixed line.

10.43. Determine the points of greatest curvature (i) on the curve $6xy = x^4+3$; (ii) on the curve $x = te^{-2t}$, $y = t^2e^{-2t}$.

10.44. Show that the radius of curvature of the curve $ay^2 = x(a-x)^2$ at the origin is $\frac{1}{2}a$, and find the radius of curvature of each branch of the curve at the point $(a, 0)$.

10.5. The perpendicular from the origin O meets the tangent at a point P of the curve in Q; r, p, and t are the lengths of OP, OQ, and QP, and ϕ is the angle between OP and PQ. Prove $p = r\sin\phi$, $t = |dp/d\psi|$.

10.51. The centre of curvature at a point P of a parabola is C and the normal PC cuts the axis of the curve in G. Prove that PG is to PC as the semi-latus rectum to the distance of P from the focus.

10.52. Find the maximum and minimum values, as a varies, of the total length of the spiral

$$x = a(a-3)e^{-at}\cos t, \qquad y = a(a-3)e^{-at}\sin t, \quad t \geqslant 0.$$

Show further that the curvature at the point $t = 1/a$ is maximum for a value of a between $1\frac{1}{2}$ and 2.

10.6. Prove that the formula for the radius of curvature is invariant under a transformation of rectangular axes.

10.61. If $x = a\cos\theta+c\cos\phi$, $y = a\sin\theta+c\cos\phi$, where θ, ϕ are functions of a variable t, connected with a third function ψ by the relation $r\dot{\psi} = a\dot{\theta}\cos(\theta-\phi)+b\dot{\phi}$, and if, as t varies from t_0 to t_1, (x, y) describes a simple closed curve, and θ, ϕ both return to their original values, prove that the area enclosed by the curve is proportional to the change in ψ.

10.7. In the catenary $y = c\operatorname{ch} x/c$ show that the length of the arc measured from the vertex $(0, c)$ to any point (x, y) of the curve is $s = c\operatorname{sh} x/c$.

If the arc of the catenary joining the points $(-\xi, \eta)$, (ξ, η) together with the bounding ordinates is rotated about the x-axis, show that the volume enclosed is $\pi c(c\xi + s\eta)$.

10.71. Find the point of maximum curvature on $y = e^x$ and find the curvature at this point.

10.72. Circles are drawn through two fixed points B, C; parabolas are drawn with the perpendicular bisector of BC for axis and a fixed point of this line for vertex. Show that a point at which one of the circles cuts one of the parabolas at right angles lies either on a fixed ellipse or on the perpendicular bisector of BC, and explain why this line is a part of the locus.

10.73. The tangent and normal at a variable point P of a plane curve cut the x-axis in T and G, and ψ is the angle which the tangent makes with this axis. If the radius of curvature at P equals the distance between T and G, prove that

$$\tan^2\tfrac{1}{2}\psi = \{(y-a)/(y+a)\}^2$$

and that the equation of the curve has one of the two forms

$$y^2 = a^2(1+be^{x/a}), \qquad y^2 = a^2(1-be^{-x/a}).$$

10.74. Find the curvature of the curve $y = \log\sec x$ at any point P of the curve, and show that the length of the arc from the origin to P is $\log(\sec x + \tan x)$, $-\frac{1}{2}\pi < x < \frac{1}{2}\pi$.

10.75. A curve C touches the x-axis at the origin and s is the length of the arc of the curve from the origin to a point (x, y) on C. The tangent to C at (x, y) meets the line $x = a$ at a distance Y from the x-axis. If dY/ds is constant, find the equation of the curve.

10.76. Q is any point on a circle, of which OA is a diameter. OQ, produced, meets the tangent to the circle at A in a point T, and the circle centre O, radius QT, meets OT at P. Prove that the area bounded by the curve on which P lies and the tangent to the circle at A is three times the area of the circle.

10.77. Show that the curvature at the point (h, k) on the curve

$$y-k = a_0(x-h)+a_1(x-h)^2/2!+a_2(x-h)^3/3!+\ldots$$

is $a_1/(1+a_0^2)^{\frac{3}{2}}$.

10.78. Find the curvatures of the curves

$$y = x(x-2)/(x-4), \qquad y^2 = x^2(x-2)/(x-4)$$

at the point $(0, 0)$.

10.79. Circles are drawn through the two points $(-1, 0)$, $(3, 0)$ and in each circle the diameter is chosen which passes through the origin. Show that the ends of these diameters lie on the curve

$$(x-2)y^2+x(x+1)(x-3) = 0.$$

10.791. Find the curvature of the curve $(x^2+y^2)^2 = a^2(y^2-x^2)$ at the point $(0, a)$, $a > 0$.

10.792. Show that the centres of curvature at points of a cycloid lie on an equal cycloid.

10.793. Find the curves on which lie the points where a curve of the family $x^2+y^2 = c$ (i) cuts at right angles, (ii) touches, a circle whose centre is the origin of coordinates.

10.794. Show that the radius of curvature of the curve

$$x = 2t-\tan t, \qquad y = 2\log\sec t$$

is $\frac{1}{2}(1+s^2)$, where s is the length of the arc measured from the origin of coordinates.

10.795. Find the area of the surface formed by rotating about the y-axis the arc of the curve $8y = x^4+2/x^2$ from $x = 1$ to $x = 2$.

10.8. If s is the arc length of the curve $\mathbf{x = x(s), y = y(s)}$ then the point $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$, where $\bar{\mathbf{x}} = \int_a^b \mathbf{x}\, ds \Big/ \int_a^b ds$, $\bar{\mathbf{y}} = \int_a^b \mathbf{y}\, ds \Big/ \int_a^b ds$, is called the *centroid* of the arc from s = a to s = b.

Find the centroid of the arc of the cycloid

$$x = a(2\psi+\sin 2\psi), \qquad y = a(1-\cos 2\psi)$$

from $\psi = 0$ to $\psi = \frac{1}{2}\pi$.

10.81. The point $(\bar{x}, \bar{y})$, where

$$\bar{x} = \int_a^b x\mu(s)\, ds \Big/ \int_a^b \mu(s)\, ds, \qquad \bar{y} = \int_a^b y\mu(s)\, ds \Big/ \int_a^b \mu(s)\, ds,$$

is called the *weighted centroid* of the arc of the curve $x = x(s)$, $y = y(s)$ from $s = a$ to $s = b$, with associated density function $\mu(s)$, or briefly, the centroid of the 'wire' $x = x(s)$, $y = y(s)$ of density $\mu(s)$.

A 'wire' in the form of the catenary $y = \operatorname{ch} x$ has density inversely proportional to y^3. Show that if A is the vertex $(0, 1)$, and P any point of the curve, the centroid of the arc AP is the point

$$x-(y/s)\log y, \qquad (y/s)\tan^{-1}s,$$

and that as s increases the centroid approaches the point $(\log 2, \frac{1}{4}\pi)$.

10.811. The centroid of the area bounded by the curves $y = y_1(x)$ and $y = y_2(x)$ and the lines $x = a$, $x = b$ is the point

$$\int_a^b x(y_1-y_2)\, dx \Big/ \int_a^b (y_1-y_2)\, dx, \qquad \frac{1}{2}\int_a^b (y_1^2-y_2^2)\, dx \Big/ \int_a^b (y_1-y_2)\, dx.$$

Find the centroid of the area bounded by the x-axis and the curve $y = \sin x$ from $x = 0$ to $x = \frac{1}{2}\pi$.

10.812. The weighted centroid of the area bounded by $y = y_1(x)$ and $y = y_2(x)$ and $x = a$, $x = b$, with associated density $\mu(x)$, or briefly the centroid of a 'lamina' of density $\mu(x)$, is the point

$$\int_a^b x(y_1-y_2)\mu(x)\, dx \Big/ \int_a^b (y_1-y_2)\mu(x)\, dx,$$

$$\frac{1}{2}\int_a^b (y_1^2-y_2^2)\mu(x)\, dx \Big/ \int_a^b (y_1-y_2)\mu(x)\, dx$$

Show that if the density of a circular lamina at any point is proportional to the distance of the point from a fixed tangent at the circumference, the centroid is distant from the centre one-quarter of the radius.

10.82. Prove that the area of a surface of revolution is equal to the product of the length of the arc revolved into the length of the path of the centroid of the arc.

10.83. Prove that the volume enclosed by the revolution of an arc about an axis is equal to the product of the area bounded by the arc and the axis into the length of the path of the centroid of that area.

10.84. The centroid of the surface formed by revolving the arc of the curve $y = f(x)$ from $x = a$ to $x = b$, about the x-axis, is the point

$$\int\limits_a^b 2\pi yx\frac{ds}{dx}\,dx\Big/\int\limits_a^b 2\pi y\frac{ds}{dx}\,dx,\quad 0,$$

and the centroid of the volume enclosed by this surface is the point

$$\int\limits_a^b \pi y^2x\,dx\Big/\int\limits_a^b \pi y^2\,dx,\quad 0.$$

Show that the x-coordinate of the centroid of the surface formed by revolving the arc of the parabola $y^2 = 4ax$, from $x = 0$ to $x = 3a$, about the x-axis, is $58a/35$.

10.9. The inclination of the tangent at the point (x,y) of a curve $y = f(x)$ is ψ, the radius of curvature at this point is $|\sigma|$, and the centre of curvature is (a,b). Prove that if σ has the same sign as $f''(x)$ then

$$a = x-\sigma\sin\psi,\qquad b = y+\sigma\cos\psi.$$

10.91. Show that the length of a quadrant of the astroid $x = a\cos^3t$, $y = a\sin^3t$ is equal to the radius of the greatest circle which has contact of the second order with the curve (i.e. the greatest radius of curvature), and that the area of the surface generated by revolving the astroid about the x-axis is $12\pi a^2/5$.

10.92. Show that the formula for the area bounded by a closed curve is invariant (unchanged) by a transformation of rectangular axes.

10.93. Show that the evolute of the parabola $x = a\mu^2$, $y = 2a\mu$ is

$$27ay^2 = 4(x-2a)^3.$$

10.94. If (S,Ψ) is the evolute of the curve (s,ψ), prove that $S = \pm\dfrac{ds}{d\psi}$, $\Psi = \psi\pm\frac{1}{2}\pi$, the signs being chosen to make S positive and Ψ lie in $(-\frac{1}{2}\pi, +\frac{1}{2}\pi)$.

10.95. If $(c\,\mathrm{sh}\,\psi,\psi)$ are the polar coordinates of a point N relative to an origin O, and P, Q are the points whose polar coordinates relative to N as origin are $(c\,\mathrm{ch}\,\psi,\psi+\frac{1}{2}\pi)$ and $(c\,\mathrm{ch}\,\psi,\psi+\frac{3}{2}\pi)$, show that P describes the evolute of the curve described by Q, and that if s is the arc-length of the latter curve, measured from O, and if $r = OQ$, then $s^2 = 2(r^2-c^2)$.

10.96. Find the total length of the spiral

$$x = e^{-t}(\sin t+\cos t),\qquad y = e^{-t}(\sin t-\cos t),\quad t \geqslant 0,$$

and show that the locus of the centre of curvature is an equal spiral.

10.97. A regular polygon of n sides rolls on a straight line. Show that the length of the path described by a vertex of the polygon in a complete revolution is $4a\sum_{r=1}^{n-1}\frac{\pi}{n}\sin\frac{r\pi}{n}$, where a is the circumradius of the polygon. Deduce that if a circle of radius a rolls on a line, in a complete revolution, the length of the path of a point on the circumference is $8a$.

10.98. P_1P_2 is a variable chord of a curve S and P_1P_2 touches a curve S^* at O. Show that when the length of the chord P_1P_2 is a maximum or minimum the normals to S at P_1, P_2 and the normal to S^* at O, are concurrent.

XI

11. Solve the following differential equations:

$(D^3-D^2-D+1)y = x(e^{-x}+1).$

$(D-1)^3y = 2\operatorname{sh}x.$

$(D^2-2D+1)y = x^2e^x.$

$(D^2+2D-3)y = x^2.$

$(D^2-2D+1)y = x^4.$

$(D^2+1)^2y = e^x+3x^2.$

$(D-1)^2y = e^x\sec^2x.$

$(D+1)^4y = e^{-x}+2x^3.$

$(D^2+4)^3y = \cos 2x.$

$(D^4+D^2+1)^2y = \sin x.$

$(D^2-4D+4)y = e^x(x^2+\sin x).$

$(D^2-4D+3)y = x^3+10\sin x.$

$(D^2-2D+1)y = xe^{3x}+\sin 2x.$

$(D^2+D+1)y = 7xe^{2x}+\sin 3x.$

$(D^2-5D+6)y = 2x\operatorname{ch}x.$

$(D^2+2D-3)y = 2\cos^2x.$

$(D^3-D^2-D+1)y = 2x^2(x+\operatorname{sh}x).$

$(D^2+4D-5)y = 1+\sin x\cos x.$

$(D^2+2D+5)^2y = e^{-x}+\sin x.$

11.1. Solve the equations:

(i) $(D-2)^2y = xe^x(1+\cos x)$; (ii) $(D^4+D^2+1)y = x\sin x$;

(iii) $(D^2-6D+13)^2y = e^{3x}(\sin x+\sin 2x)$.

11.2. Solve the equation $x\,dy/dx+x = (x+y)\log(x+y)$ by means of the substitution $z = x+y$.

11.21. Solve the equation $x^2\,d^2y/dx^2 = (2x^2+3)(x\,dy/dx-y)$ by the substitution $z = x\,dy/dx-y$.

11.3. Show that the substitution $x = e^t$ transforms the equation $x^2(d^2y/dx^2)-x(dy/dx)+y = x$ into $d^2y/dt^2-2dy/dt+y = e^t$ and hence solve the equation.

11.31. Transform the equation $dy/dx = 2+3y+y^2$ by the substitution $z = e^{-\int y\,dx}$, and hence solve the equation.

11.4. Find the solution of the equation $(d^2y/dx^2)^2 = 4(dy/dx)$ which is such that at $x = 0$, $y = 1$ and $dy/dx = 0$.

11.41. Find the solution of the equation $d^2y/dx^2 = 2\sin 2y$ which is such that at $x = 0$, $y = \frac{1}{2}\pi$ and $dy/dx = 2$.

11.5. Verify that the differential equation $d^2y/dx^2 = 2(1+2x^2)y$ has a solution of the form $y = e^{x^n}$, and hence solve the equation completely.

11.6. Prove that

$$1+\frac{x^3}{3!}+\frac{x^6}{6!}+\frac{x^9}{9!}+\dots = \tfrac{1}{3}e^x+\tfrac{2}{3}e^{-\frac{1}{2}x}\cos\frac{x\sqrt{3}}{2}.$$

11.7. If $\frac{d}{dx}x^y = kyx^{y-1}$, prove that $y = a(\log x)^{k-1}$ where a, k are constants.

11.8. If $L\left\{\frac{1}{D\phi(x)}\right\} = \frac{1}{\phi(x)}$, prove that $\phi(x) = \pm\sqrt{(ax+b)}$, where a and b are constant.

11.9. If $t_1 = e^{px}$, $t_2 = xe^{px}$, $u_1 = \sin ax$, $u_2 = \cos ax$, $v_1 = x\sin ax$, and $v_2 = x\cos ax$, prove that the Wronskians of t_1, t_2, of u_1, u_2, and of u_1, u_2, v_1, v_2 do not vanish for any value of x.

11.91. If $\mathbf{L(t)}$ is a polynomial of the nth degree, prove that there is only one solution η of the equation $\mathbf{L(D)y = 0}$ such that η, $\mathbf{D}\eta$, $\mathbf{D}^2\eta,\ldots, \mathbf{D}^{n-1}\eta$ take assigned values at a given point $\mathbf{x = a}$.

In particular, show that if η, $\mathbf{D}\eta,\ldots, \mathbf{D}^{n-1}\eta$ are all zero at $\mathbf{x = a}$, then $\eta = 0$ for all values of $\mathbf{x}$.

11.92. The functions $\eta_1, \eta_2,\ldots, \eta_n$ are solutions of the nth order equation $\mathbf{L(D)y = 0}$. Prove that, if $\mathbf{W}(\eta_1, \eta_2,\ldots, \eta_n)$, the Wronskian of $\eta_1, \eta_2,\ldots, \eta_n$, vanishes at the point $\mathbf{x = a}$, then it is zero for all values of $\mathbf{x}$.

11.93. Find the complete solution of the simultaneous equations

$$(D^2-1)u+(D+4)v = 2x-x^2+\sin x,$$
$$(D+1)u-Dv = x^2.$$

11.94. If $\phi(x)$ stands for either $\sin x$ or $\cos x$, prove that

$$(D^2+a^2)^n x^{n+p}\phi(ax)$$
$$= (n+p)!\left\{\frac{x^p}{p!}(2D)^n+\frac{x^{p-1}}{(p-1)!}n(2D)^{n-1}+\frac{x^{p-2}}{(p-2)!}\binom{n}{2}(2D)^{n-2}+\ldots\right\}\phi(ax).$$

XII

12. If $f(x)$ has n roots in (a,b), prove that $f^{(n-1)}(x)$ has at least one root in that interval.

12.01. If $a < b < c$, $a+b+c = 2$, $ab+bc+ca = 1$, prove that a, b, c lie in the intervals $[0,\frac{1}{3}]$, $[\frac{1}{3},1]$, $[1,\frac{4}{3}]$ respectively.

12.1. If $f(x)$ is differentiable in (a,b) and $f(\lambda) = 0$, $a < \lambda < b$, prove that there is a continuous function $g(x)$ such that $f(x) = (x-\lambda)g(x)$.

12.11. If c_x is the function, given by the mean-value theorem, such that

$$\{f(x)-f(a)\}/(x-a) = f'(c_x), \quad a < x \leqslant b$$

and $c_a = a$, prove that c_x is continuous in (a,b), provided $|f''(x)| > \alpha > 0$ in (a,b).

12.12. If $g(a) = g(b) = 0$ and $g''(x) \neq 0$ in $[a,b]$, prove that $g(x) \neq 0$ in (a,b).

12.2. If $\nu(x,y)$ denotes $\{f(y)-f(x)\}/\{g(y)-g(x)\}$, where $g(x)$ is monotonic increasing, prove that if $\nu(a,X) > \nu(a,b)$, $a < X < b$, then $\nu(X,b) < \nu(a,b)$, and if $\nu(a,X) < \nu(a,b)$ then $\nu(X,b) > \nu(a,b)$.

12.21. If $f(x)$ and $g(x)$ are differentiable in (a,b), and $g'(x) \neq 0$ in (a,b), show that if $\nu(a,x)$ is constant in (a,b) then $\nu(a,b) = f'(c)/g'(c)$ for any c in (a,b).

12.22. If $f(x)$ and $g(x)$ are differentiable in (a,b), and $g'(x) \geqslant 2\lambda > 0$ in (a,b), prove that there is a c_1 in (a,b) such that $f'(c_1)/g'(c_1) \leqslant \nu(a,b)$ and a c_2 such that $f'(c_2)/g'(c_2) \geqslant \nu(a,b)$.

12.23. Deduce from 12.22 that if $\nu(a,x)$ is not constant in (a,b) then there is a point c in (a,b) such that $\nu(a,b) = f'(c)/g'(c)$.

12.3. Prove there is a point α in $(-h,h)$ such that

$$\int_{-h}^{h} f(x)\,dx = 2hf(0)+(h^3/3)f''(\alpha).$$

12.31. Prove there is a point β in $(-h,h)$ such that

$$\int_{-h}^{h} f(x)\,dx = h\{f(h)+f(-h)\}-(2h^3/3)f''(\beta).$$

12.32. Prove there is a point γ in $(-h,h)$ such that

$$\int_{-h}^{h} f(x)\,dx = (h/2)\{f(h)+2f(0)+f(-h)\}-(h^3/6)f''(\gamma).$$

12.33. Prove there is a point δ in $(0,h)$ such that

$$\int_{0}^{h} f(x)\,dx = (h/12)[-f(-h)+8f(0)+5f(h)]-(h^4/4!)f'''(\delta).$$

12.331. If $f(x)$ is differentiable three times in $(a-h,a+h)$, prove there is a point c in this interval such that

$$\frac{f(a+h)-f(a-h)}{2h}-f'(a) = \frac{h^2}{6}f'''(c).$$

12.332. If $g(x)$ is differentiable twice in $(a-2h,\ a+2h)$, prove there is a point c in the interval such that

$$g(a+2h)-g(a+h)-g(a-h)+g(a-2h) = 3h^2g''(c).$$

12.333. If $f(x)$ is differentiable five times in $(a-2h,a+2h)$ prove there is a point c in the interval such that

$$\frac{f(a+2h)-8f(a+h)+8f(a-h)-f(a-2h)}{12h}+f'(a) = \frac{h^4}{30}f^5(c).$$

12.334. If $f(x)$ is differentiable four times in $(a-h,a+h)$ prove there is a point c in the interval such that

$$\frac{f(a+h)-2f(a)+f(a-h)}{h^2}-f''(a) = \frac{h^2}{12}f^4(c).$$

12.335. If $f(x)$ is differentiable six times in $(a-2h,a+2h)$ prove there is a point c, $a-2h < c < a+2h$, such that

$$\frac{f(a+2h)-16f(a+h)+30f(a)-16f(a-h)+f(a-2h)}{12h^2}+f''(a) = \frac{h^4}{90}f^6(c).$$

12.4. If $f(x) > 0$ and $g(x)$ is monotonic and bounded, for all values of x, and if $\int^{\infty} f(x)\,dx$ exists then $\int^{\infty} f(x)g(x)\,dx$ exists.

12.41. If $g(x)$ is monotonic decreasing to zero and $\left|\int^{x} f(x)\,dx\right| < M$ for all values of x then $\int_a^\infty f(x)g(x)\,dx$ exists.

12.5. If $f(x)$ is continuous in any interval (h, H), $0 < h < H$, and if $f(a_n) \to \lambda$, and $f(1/a_n) \to \mu$, when $a_n \to 0$, then

$$\int_0^\infty \{f(ax)-f(bx)\}\,dx/x = (\lambda-\mu)\log(b/a), \quad b > a > 0,$$

12.51. If $\int_a^b f(Nx)\,dx/x \to 0$ as N increases, and $f(a_n) \to \lambda$ when $a_n \to 0$, then

$$\int_0^\infty \{f(ax)-f(bx)\}\,dx/x = \lambda\log(b/a).$$

12.52. If $\int_a^b f(x/n)\,dx/x \to 0$ and $f(1/a_n) \to \mu$ when $a_n \to 0$ then

$$\int_0^\infty \{f(ax)-f(bx)\}\,dx/x = \mu\log(a/b).$$

12.6. Prove that if $0 < X < \pi$ then $\int_0^X \{D_x(1/\sin x - 1/x)\}\cos ax\,dx$ exists and has an absolute value less than $1/\sin X - 1/X$.

12.61. Prove that
$$\int_0^{\frac{1}{2}\pi} \frac{\sin(2n+1)x}{\sin x}\,dx = \tfrac{1}{2}\pi,$$

and deduce that
$$\int^\infty \frac{\sin x}{x}\,dx = \tfrac{1}{2}\pi.$$

12.62. Prove that

$$\int^\infty \frac{1-\cos x}{x^2}\,dx = \int^\infty \left(\frac{\sin x}{x}\right)^2 dx = \tfrac{1}{2}\pi.$$

12.63. Prove that
$$\begin{aligned}\int^\infty \frac{\sin ax}{x}\,dx &= \tfrac{1}{2}\pi, & a > 0,\\ &= 0, & a = 0,\\ &= -\tfrac{1}{2}\pi, & a < 0.\end{aligned}$$

12.64. Prove $\left(1+\dfrac{1}{x}\right)^x$ increases for $x > 0$ and $\left(1-\dfrac{1}{x}\right)^{-x}$ decreases for $x > 1$. Hence show that $\left(1+\dfrac{1}{x}\right)^x$ and $\left(1-\dfrac{1}{x}\right)^{-x}$ tend to a common limit as $x \to \infty$.

12.7. Prove that, if $|f^{\mathrm{vi}}(x)| < M$ in $(-3h, 3h)$, then

$$\int_{-3h}^{3h} f(x)\,dx = \tfrac{3}{10}h[f(-3h)+5f(-2h)+f(-h)+6f(0)+f(h)+5f(2h)+f(3h)]$$

with an error less than $\frac{3}{5}h^7M$.

12.8. Illustrate the necessity for the condition '$g(x) > 0$' in Theorem 12.61 by showing that if $f(x) = (x-1)^{\frac{1}{3}}$, $g(x) = (x-1)$ then

$$f(X)\int_0^3 g(x)\,dx < \int_0^3 f(x)g(x)\,dx$$

for any X in $(0, 3)$.

12.81. Illustrate the necessity for the condition '$f(x)$ is of constant sign' in Theorem 12.73 by showing that if $f(x) = 1-x$, $g(x) = \sqrt{x}$ then

$$\int_0^4 f(x)g(x)\,dx < f(0)\int_0^X g(x)\,dx$$

for any X in $(0, 4)$.

12.82. If $f(0) = a$, $f(a) = b$, $f'(0) = -1$, and if $|f''(x)| < 1/4|a|$ in the interval $(-2a, 2a)$ prove that $|1+f'(x)| < \frac{1}{2}$ in $(-2a, 2a)$ and hence that

$$|f(a+b)| < \tfrac{1}{2}|f(a)| < \tfrac{1}{4}|a|.$$

XIII

13. Prove that

$$\tan x = x+\tfrac{1}{3}x^3+x^5(\tfrac{2}{15}+\alpha_x), \qquad \sin^{-1}x = x+\tfrac{1}{6}x^3+x^5(\tfrac{3}{40}+\beta_x),$$

where α_x and β_x both tend to zero with x.

13.01. Evaluate the limits:

$$\lim_{x\to 0}(1/\sin x - 1/x), \qquad \lim_{x\to 1}(1+\cos\pi x)/\tan^2\pi x,$$

$$\lim_{x\to 0}\{\log(1+x)\}^3/(\tan x - \sin x), \qquad \lim_{x\to\frac{1}{2}\pi}\cos 3x/(e^{2x}-e^{\pi}),$$

$$\lim_{x\to 0}(\sin x\sin^{-1}x - x^2)/(\tan x\tan^{-1}x - x^2), \qquad \lim\{\cot^2 x-(x-\pi)^{-2}\},$$

$$\lim_{x\to 0}\{1/x^4 - 3\sin x/x^5(2+\cos x)\}.$$

13.02. Prove that $\qquad \lim_{n\to\infty} n\{\sqrt[n]{x}-1\} = \log x.$

13.03. Prove that e^x exceeds or is exceeded by $1+x+\dfrac{x^2}{2!}+\dots+\dfrac{x^n}{n!}$ according as x^{n+1} is positive or negative.

13.1. Given a function $F(x)$ and a function $G(x)$ such that $G(a) = 0$, $G(b) = 1$, find constants A, B such that the function $F(x)+A+BG(x)$ is zero when $x = a$ and when $x = b$. Taking

$$F(x) = f(b)-f(x)-\frac{(b-x)}{1!}f'(x)-\frac{(b-x)^2}{2!}f^2(x)-\dots-\frac{(b-x)^n}{n!}f^n(x),$$

where $f(x)$ has $n+1$ successive derivatives, choose $G(x)$ and deduce a form of Taylor's theorem for the function $f(x)$.

13.11. In 13.1 take $G(x)$ such that $G(0) = 1$, $G(h) = 0$, $h = b-a$; determine A, B so that $F(x)+A+BG(x)$ is zero when $x = 0$ and when $x = h$ and take

$$F(x) = f(b)-f(b-x)-xf'(b-x)-\frac{x^2}{2!}f^2(b-x)-\ldots-\frac{x^n}{n!}f^n(b-x).$$

13.12. In 13.1 take $G(0) = 0$, $G(h) = 1$; determine A and B so that

$$F(x)+A+BG(x)$$

is zero when $x = 0$ and when $x = h$ and take

$$F(x) = f(b)-f(a+x)-(h-x)f'(a+x)-\ldots-\frac{(h-x)^n}{n!}f^n(a+x).$$

13.13. If h and k have the same sign, prove that there is a λ such that

$$f(a+h)+f(a+k) = 2f(a)+(h+k)f'(a)+$$
$$+\left(\frac{h^2+k^2}{2!}\right)f''(a)+\ldots+\left(\frac{h^{n-1}+k^{n-1}}{(n-1)!}\right)f^{n-1}(a)+\frac{h^n+k^n}{n!}f^n(\lambda).$$

13.14. Prove that, for any h and k we can find θ, $0 < \theta < 1$, such that

$$f(a+h)-f(a+k) = (h-k)f'(a)+$$
$$+\left(\frac{h^2-k^2}{2!}\right)f''(a)+\ldots+\left(\frac{h^{n-1}-k^{n-1}}{(n-1)!}\right)f^{n-1}(a)+\frac{h^nf^n(a+\theta h)-k^nf^n(a+\theta k)}{n!}.$$

13.2. If θ_n is given by Taylor's theorem so that

$$f(a+h) = f(a)+hf'(a)+\frac{h^2}{2!}f''(a)+\ldots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(a)+\frac{h^n}{n!}f^n(a+\theta_n h),$$

prove that, provided $f^{n+1}(a) \neq 0$, then $\theta_n \to 1/(n+1)$.

13.21. If $f(x)$ is continuous in the interval (α, b) for any α greater than a, and if $f(x) \to f(a)$ as $x \to a$, prove that $f(x)$ is continuous in (a, b).

13.3. Find the root of the equation $x+e^x = 9.3$ correct to 2 decimal places, given that $x = 2$ is an approximate solution.

13.31. If θ is small show that an approximation to a root of the equation $\sin x = \theta x$ is $\pi\{1-\theta+\theta^2-(1+\pi^2/6)\theta^3\}$.

13.32. Find the positive (non-zero) root of the equation $3\sin x = 2x$ correct to 2 decimal places.

13.4. If $y^3+y-x = 0$ show that $y = x-x^3+3x^5(1+\epsilon_x)$, where $\epsilon_x \to 0$ as $x \to 0$.

13.41. If $y = e^{xy}$ show that $y = 1+x+\frac{3}{2}x^2+\frac{8}{3}x^3(1+\epsilon_x)$, where $\epsilon_x \to 0$ as $x \to 0$.

13.5. Show that the conic of closest contact with $y = ax^2+bx^3+cx^4$ at the origin is

$$a^3y = a^4x^2+a^2bxy+(ac-b^2)y^2.$$

13.51. Taking the tangent and normal at a point P of a curve as axes show that the parabola of closest contact at P is $(y\,d\rho/ds-3x)^2 = 18\rho y$, where ρ and $d\rho/ds$ are the values of the radius of curvature and its derivative at the point P of the curve.

13.6. If ρ and ρ' are the values of the radius of curvature and its

derivative at the origin, show that (in the neighbourhood of the origin) the equation of a curve touching the x-axis at the origin is

$$x = s - s^3/6\rho^2 + \ldots, \qquad y = s^2/2\rho - \rho' s^3/6\rho^2 + \ldots.$$

13.601. Prove that the sequences $\sqrt{n}\,x(1-x)^n$ and $x^n(1-x)$ are both interval-convergent in $(0,1)$, and that the sequence $nx(1-x)^n$ is not interval-convergent in $(0,1)$.

13.61. Prove that

$$e^{x\cos\alpha}\sin(x\sin\alpha) = x\sin\alpha + \frac{x^2}{2!}\sin 2\alpha + \frac{x^3}{3!}\sin 3\alpha + \ldots,$$

$$e^{x\cos\alpha}\cos(x\sin\alpha) = 1 + x\cos\alpha + \frac{x^2}{2!}\cos 2\alpha + \frac{x^3}{3!}\cos 3\alpha + \ldots.$$

13.62. If $|a_n(x)| \leqslant u_n$, $a \leqslant x \leqslant b$, and $\sum u_n$ is convergent, prove that $\sum a_n(x)$ is interval-convergent in (a,b).

13.63. If $\phi(x)$ is differentiable n times, prove

$$\int\limits_0^c \frac{\phi(x)}{x^{\frac{1}{2}}(c-x)^{\frac{1}{2}}}\,dx = a_0 + a_1 c + a_2 c^2 + \ldots + a_{n-1}c^{n-1} + c^n(a_n + \epsilon_c),$$

where $r!\,a_r = 2\phi^r(0)\int\limits_0^{\frac{1}{2}\pi}\sin^{2r}\theta\,d\theta$ and $\epsilon_c \to 0$ as $c \to 0$.

13.64. The curves $y = f(x)$, $y = g(x)$ have contact of the nth order exactly at the point $x = a$; prove that the curves cross, or do not cross, at $x = a$ according as n is even or odd.

13.7. If $\sum v_r(x)$ is interval-convergent in the interval $(0,\infty]$, and if for each value of r, $v_r(x) \to w_r$ as $1/x \to 0$, prove that $\sum w_r$ converges and that if $1/p_n \to 0$ then

$$\sum_{r=0}^{p_n} v_r(n) \to \sum w_r.$$

13.71. If, for each value of r, $v_r(n) \to w_r$ and $\sum |v_r(x)|$ is interval-convergent in the interval $(0,\infty]$, and if $1/p_n \to 0$, prove that

$$\prod_1^{p_n}\{1+v_r(n)\} \to \prod_{r\geqslant 1}(1+w_r),$$

where $\prod\limits_1^k a_r$ denotes $a_1 a_2 a_3 \ldots a_k$ and $\prod\limits_{r\geqslant 1} a_r$ denotes $\lim a_1 a_2 \ldots a_n$.

13.72. Prove that

$$\frac{\sin x}{x} = \left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{2^2\pi^2}\right)\left(1-\frac{x^2}{3^2\pi^2}\right)\left(1-\frac{x^2}{4^2\pi^2}\right)\ldots.$$

13.8. If $f(x)$ is differentiable twice in $(a, a+h)$ and $0 < r < 1$, show that the error in taking $f(a)+r\{f(a+h)-f(a)\}$ for $f(a+rh)$ is $\frac{1}{2}r(1-r)h^2 f''(a+\theta h)$ for a certain θ in $(0,1)$.

Hence, considering $\sin x$ in $(\frac{1}{6}\pi, \frac{1}{4}\pi)$ prove $\sin\dfrac{7\pi}{36} = \dfrac{2+\sqrt{2}}{6}$ with an error less than $\pi^2\sqrt{2}/2592$.

13.81. If $0 \leqslant x \leqslant a-\delta$, prove $|(a-x)^{1/n}-1| < \delta$ for $n > n_0$, where n_0 is independent of x.

If $f(x)$ is continuous in $(0,a)$, and $f(a) = 0$, and if $f_n(x) = (a-x)^{1/n}f(x)$, prove that $f_n(x)$ is interval-convergent to $f(x)$ in $(0,a)$.

If further $f(x)$ has a derivative $f'(x)$ in $(0,a)$, and if $f'(a) = 0$, prove that $f_n'(x)$ is interval-convergent to $f'(x)$ in $(0,a)$.

13.82. If $y = e^{-x}\left(1+\frac{x}{n}\right)^n$ prove that there is a point c in $(0,x)$ such that

$$y = 1-\frac{x^2}{2n}+\frac{x^3}{3!}R(c),$$

where
$$R(c) = \frac{1}{n^3}\left\{(2n+3nc-c^3)e^{-c}\left(1+\frac{c}{n}\right)^{n-3}\right\},$$

and deduce that
$$n\left[1-e^{-x}\left(1+\frac{x}{n}\right)^n\right] \to \frac{x^2}{2}.$$

13.83. (i) If $f'(x) \to l > 0$, as $x \to \infty$, prove that $\frac{f(x)}{x} \to l$.

(ii) If $\theta'(x) \geqslant \lambda > 0$ when $x \geqslant 0$, and $\theta'(x) \to l$ as $x \to \infty$, and if

$$\phi(x) = \int_0^x \{1/\theta'(t)\}\,dt, \qquad \psi(x) = \theta(x)\phi(x)$$

prove that $\frac{\psi(x)}{x^2} \to 1$, as $x \to \infty$.

13.9. The function $f(x)$ is said to be *point-differentiable* at a point $x = c$ if there is a number ϕ such that

$$\frac{f(x)-f(c)}{x-c} \to \phi \quad \text{as} \quad x \to c.$$

The number ϕ is called the *point-derivative* of $f(x)$ at $x = c$.

If $f(x)$ is point-differentiable at $x = c$, with point-derivative ϕ, and if $a_n < c < b_n$ and $b_n - a_n \to 0$, prove that

$$\frac{f(b_n)-f(a_n)}{b_n-a_n} \to \phi.$$

13.91. If $f(x)$ is point-differentiable at each point x of an interval (a, b), with point-derivative $\phi(x)$, and *if* $\phi(x)$ is continuous in (a, b), prove that for any two points α, β in (a, b) there is a point γ between α and β such that

$$f(\beta)-f(\alpha) = (\beta-\alpha)f'(\gamma).$$

13.92. If $f(x)$ is point-differentiable at each point x of an interval (a, b), with point-derivative $\phi(x)$, and *if* $\phi(x)$ is continuous in (a, b), prove that $f(x)$ is interval-differentiable in (a, b), with interval-derivative $\phi(x)$.

XIV

14. Prove that

14.01, $\sum_{r=1}^{\infty} n/(n^2+r^2) \to \frac{1}{4}\pi$. .02, $\sum_{r=1}^{4n} 1/\sqrt{(rn)} \to 4$. .03, $\sum_{r=2n}^{3n} 2r/(n^2+r^2) \to \log 2$.

14.1. Prove that

$$\sum_{r=1}^{\infty} 2rn^2/(n^4+r^2n^2+r^4) \to \pi/3\sqrt{3}.$$

14.2. Show that $\int_0^{\frac{1}{2}} e^{x^2}\,dx = \cdot 545$ with an error less than $4/10^4$.

14.21. Evaluate $\int_1^3 \frac{\log x}{x}\,dx$ correct to 3 decimal places.

14.22. Evaluate $\int_{\frac{1}{2}\pi}^{\pi} \frac{\sin x}{x}$ to 5 decimal places.

14.3. Prove that

$$\sum \frac{\sin nx}{n^3} = \frac{1}{12}x(x-\pi)(x-2\pi)$$

and
$$\sum \frac{\cos nx}{n^4} = \frac{\pi^4}{90} - \frac{1}{48}x^2(x-2\pi)^2, \qquad 0 \leqslant x \leqslant 2\pi.$$

14.4. If $f(x)$ is positive, continuous, and steadily decreasing as x increases, prove that

$$f(n)+f(n+1)+...+f(N) - \int_n^N f(x)\,dx$$

lies between zero and $f(n)$; hence show that if $f(n) \to 0$ then $\sum_1^n f(r) - \int_1^n f(x)\,dx$ is convergent.

14.41. Show that $1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n}-\log n$ converges to a limit between zero and unity (this limit is known as Euler's constant, and denoted by γ).

14.42. If $f(x)$ is positive, continuous, and steadily decreasing as x increases, prove that $\sum_1^n f(x)$ and $\int_1^n f(x)\,dx$ converge and diverge together.

14.5. The *arithmetic mean* $A_{a,b}^f$ of a function $f(x)$ over an interval (a, b) is defined to equal $\left\{\int_a^b f(x)\,dx\right\}\Big/(b-a)$, and the geometric mean of $f(x)$ over (a, b) is defined to be $G_{a,b}^f = e^{A_{a,b}^{\log f}}$, so that $\log G_{a,b}^f = A_{a,b}^{\log f}$.

Evaluate the arithmetic and geometric means of x^n over (a, b).

14.501. Find the values of the arithmetic and geometric means of $\sin x$ over $(0, \frac{1}{2}\pi)$.

14.51. If $\phi(x) = f\{a(1-x)+bx\}$ prove that $A_{a,b}^f = A_{0,1}^\phi$ and $G_{a,b}^f = G_{0,1}^\phi$.

14.52. If $f(x)$ is continuous in $(0, 1)$ prove that $A_{0,1}^f$ is the limit of the arithmetic mean of the n numbers $f(0)$, $f(1/n)$, $f(2/n)$,..., $f\{(n-1)/n\}$, and that $G_{0,1}^f$ is the limit of their geometric mean.

14.6. If $f(x)$ is continuous in (a, b), $f(x) \geqslant 0$ throughout (a, b), and $f(c) > 0$ for at least one point c in (a, b), prove that $\int_a^b f(x)\,dx > 0$.

14.7. If $x_{n+1} = +\sqrt{x_n}$, $x_0 > 0$, prove that

$$\int_1^{x_{n+1}} (1/y)\,dy = \tfrac{1}{2}\int_1^{x_n} (1/x)\,dx.$$

Hence show that

$$2^n(x_n-1) \to \int^{x_0} (1/x)\,dx = \log x_0.$$

14.71. If $x_{n+1} = x_n/\{1+\surd(1+x_n^2)\}$, show by means of the transformation

$$y = x/\{1+\surd(1+x^2)\},$$

that

$$\int_0^{x_{n+1}} \frac{1}{1+y^2}\,dy = \frac{1}{2}\int_0^{x_n} \frac{1}{1+x^2}\,dx.$$

Hence show that

$$2^n x_n \to \int^{x_0} \frac{1}{1+x^2}\,dx = \arctan x_0.$$

14.8. Prove that a continuous function attains both its arithmetic and geometric means.

14.81. If $f(x,y)$ is continuous in $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$, for all values of b greater than a, and if the sequence $\int_a^n f(x,y)\,dx$ is interval-convergent in (c,d), then $\int_a^\infty f(x,y)\,dx$ is continuous in y in (c,d).

14.82. If $u_n(x) \geqslant 0$ and $\int_a^N \{\sum u_n(x)\}\,dx = \sum \int_a^N u_n(x)\,dx$ for all N, then if both the limits $\int_a^\infty \{\sum u_n(x)\}\,dx$, $\sum \int_a^\infty u_n(x)\,dx$ exist, they are equal.

Show further that the condition $u_n(x) \geqslant 0$ may be relaxed if, in addition to the remaining conditions, we have $\int_a^N \sum |u_n(x)|\,dx = \sum \int_a^N |u_n(x)|\,dx$ for all N, and if the limits $\int_a^\infty \sum |u_n(x)|\,dx$ and $\sum \int_a^\infty |u_n(x)|\,dx$ exist.

14.9. If $a = a_0^p, a_1^p, a_2^p,..., a_{r_p}^p = b$ is a p-chain of a function $f(x)$ which is semi-continuous in $(a,b]$ and if $S_p = \sum_{r=0}^{r_p-1} f(a_r^p)(a_{r+1}^p - a_r^p)$, prove that the sequence S_p is convergent.

14.91. If $a = a_0^p, a_1^p,..., a_{r_p}^p = b$ and $a = b_0^p, b_1^p,..., b_{\mu_p}^p = b$ are both p-chains of a semi-continuous function $f(x)$, and if $S_p^1 = \sum f(a_r^p)(a_{r+1}^p - a_r^p)$, $S_p^2 = \sum f(b_r^p)(b_{r+1}^p - b_r^p)$, prove that S_p^1 and S_p^2 tend to the same limit.

We define the integral over (a, b) of a function f(x) semi-continuous in (a, b], to be the limit of the sequence $S_p = \sum f(a_r^p)(a_{r+1}^p - a_r^p)$; the integral is denoted by $\int_a^b f(x)\,dx$.

14.92. If $f(x)$ is semi-continuous in $(a,b]$, and if $a < c < b$, prove that

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx.$$

14.93. If $f(x)$ and $g(x)$ are semi-continuous in $(a,b]$ prove that

$$\int_a^b f(x)\,dx+\int_a^b g(x)\,dx=\int_a^b [f(x)+g(x)]\,dx.$$

14.94. If $f(x)$ is semi-continuous in $(a,b]$ and if m, M are bounds of $f(x)$ in $(a,b]$, prove that

$$m(b-a)\leqslant\int_a^b f(x)\,dx\leqslant M(b-a).$$

14.941. If $f(x)$ is semi-continuous in $(a,b]$ prove that $\int_a^t f(x)\,dx$ is a continuous function of t, in (a,b).

14.95. If f(x) is positive, non-decreasing and semi-continuous, and g(x) is continuous, with a finite number of roots, in (a, b), prove that there is a point c in [a, b] such that

$$\int_a^b f(x)g(x)\,dx=f(b)\int_c^b g(x)\,dx.$$

14.96. Prove that

$$\cot x=\frac{1}{x}-2x\sum_1^\infty\frac{1}{k^2\pi^2-x^2}.$$

(Use Example 13.72.)

14.961. Prove that $\dfrac{1}{\sin x}=\dfrac{1}{x}+2x\sum_1^\infty\dfrac{(-)^k}{x^2-k^2\pi^2}.$

14.962. Prove that

$$\int_0^\infty\frac{x^{a-1}}{1+x}\,dx=\frac{\pi}{\sin a\pi},\quad 0<a<1.$$

XV

15. If $x\cos u+y\sin u=1$, $v=x\sin u-y\cos u$ prove that

$$v^2\frac{\partial^2 u}{\partial x\partial y}+v\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}=\cos 2u.$$

15.01. Transform the equation $n^2\dfrac{\partial^2 H}{\partial x^2}=\dfrac{\partial^2 H}{\partial y^2}$ by the substitution

$$u=x+ny,\qquad v=x-ny,$$

where n is a constant, and hence show that $H=f(x+ny)+g(x-ny)$, where f and g are arbitrary functions.

15.02. Prove that if $\{\phi(x)+\psi(y)\}^2e^z=2\phi'(x)\psi'(y)$ then $\dfrac{\partial^2 z}{\partial x\partial y}=e^z$.

15.03. If $\log H=\alpha t+\beta x^2/t-\frac{1}{2}\log t$, show that

$$\frac{\partial^2 H}{\partial x^2}+4\beta\frac{\partial H}{\partial t}=4\alpha\beta H.$$

15.04. If $x = r\cos\theta\sin\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\phi$, and H is a function of x, y, z, prove that

$$\cos\theta\frac{\partial H}{\partial x}+\sin\theta\frac{\partial H}{\partial y} = \sin\phi\frac{\partial H}{\partial r}+\frac{\cos\phi}{r}\frac{\partial H}{\partial\phi},$$

$$-\sin\theta\frac{\partial H}{\partial x}+\cos\theta\frac{\partial H}{\partial y} = \frac{1}{r\sin\phi}\frac{\partial H}{\partial\theta}.$$

15.05. If $f(x,y)$ is differentiable in $(a,A)(b,B)$ and $f(a,b) = f(A,B)$, prove that there is a point (α,β) on the line segment joining (a,b) to (A,B) such that

$$(\alpha-a)f_x(\alpha,\beta)+(\beta-b)f_y(\alpha,\beta) = 0.$$

15.1. Show that if z denotes either of the functions $\frac{1}{2}\log(x^2+y^2)$, $\tan^{-1}(y/x)$ then

$$\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2 = 1/(x^2+y^2).$$

15.11. If $\phi(x,y,z) = 0$, show that

$$\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y\left(\frac{\partial x}{\partial y}\right)_z = -1.$$

15.2. If $\mathbf{P(x,y)}$ is a polynomial such that $\mathbf{P(0,b) = 0,\ P_y(0,b) \neq 0}$, show that a solution of the equation $\mathbf{P(x,y) = 0}$ may be found in the form $\mathbf{y = b+a_1x+\ldots+a_{n-1}x^{n-1}+x^n(a_n+\alpha)}$, where $\boldsymbol{\alpha}$ is a differentiable function of x such that $\boldsymbol{\alpha \to 0}$ as $\mathbf{x \to 0}$.

If $\mathbf{2x^2+xy-y^2+x^3-xy^2+y^3 = 0}$, determine y as a power series in x as far as the terms in $\mathbf{x^3}$, (i) near the origin, (ii) near the point (0, 1).

15.21. If $f(x,y)$ is continuous in $(\alpha,b)(\gamma,d)$ for any α greater than a, and any γ greater than c, and if in the interval $\alpha \leqslant x \leqslant b$, $f(x,c)$ is the continuous interval limit of $f(x,y)$ as y tends to c, and in the interval $\gamma < y < d$, $f(a,y)$ is the continuous interval limit of $f(x,y)$ as x tends to a, and if $f(a,c)$ is the double limit of $f(x,y)$ as x tends to a and y tends to c, prove that $f(x,y)$ is continuous in $(a,b)(c,d)$.

15.3. If $z = e^{ax}f(x+y)+e^{-ax}g(x-y)$ show that

$$\frac{\partial^2 z}{\partial x^2} = a^2z+2a\frac{\partial z}{\partial y}+\frac{\partial^2 z}{\partial y^2}.$$

15.31. Denoting $\dfrac{\partial^2\phi}{\partial x^2}+\dfrac{\partial^2\phi}{\partial y^2}$ by $\nabla^2\phi$, for any ϕ, prove that, if $r^2 = x^2+y^2$, then $\nabla^2 r = 1/r$, $\nabla^2(1/r) = 1/r^3$, $\nabla^2\log r = 0$, and $\nabla^2\tan^{-1}y/x = 0$.

Evaluate

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)(x^2+y^2+z^2)^{\frac{1}{2}} \quad\text{and}\quad \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)(x^2+y^2+z^2)^{-\frac{1}{2}}.$$

15.4. Find the locus of the points at which a curve of the family $\boldsymbol{\phi(x,y) = \lambda}$ touches a curve of the family $\boldsymbol{\psi(x,y) = \mu}$.

A family of circles passes through the fixed points $\boldsymbol{(p,0)}$, $\boldsymbol{(q,0)}$ and a family of parabolas touches the $\boldsymbol{x}$-axis at the origin, each parabola having the origin for vertex. Show that the points where a circle of one family touches a parabola of the second, is another parabola.

15.41. Prove that the centre of curvature at the point (x, y) on the curve $\phi(x, y) = 0$ is

$$\xi = x - \phi_x(\phi_x^2+\phi_y^2)/(\phi_{xx}\phi_y^2 - 2\phi_{xy}\phi_x\phi_y + \phi_{yy}\phi_x^2),$$
$$\eta = y - \phi_y(\phi_x^2+\phi_y^2)/(\phi_{xx}\phi_y^2 - 2\phi_{xy}\phi_x\phi_y + \phi_{yy}\phi_x^2).$$

15.42. Prove that the curvature of the curve $\phi(x, y) = 0$ is k given by

$$k = \mp\{\phi_{xx}\phi_y^2 - 2\phi_{xy}\phi_x\phi_y + \phi_{yy}\phi_x^2\}/\{\phi_x^2+\phi_y^2\}^{\frac{3}{2}}.$$

Find the curvature at any point of the curve

$$\sin x + \sin y = c.$$

15.43. A variable parabola through a fixed point has a given circle of curvature at that point. Show that the focus of the parabola lies on a fixed circle, and that the directrix passes through a fixed point.

15.44. If P, P_1, P_2 are points on the curve $y = f(x)$, show that as P_1 and P_2 approach P, the centre and radius of the circle PP_1P_2 approach the centre and radius of curvature of the curve at P.

15.5. If $y = f(x, u)$ and $f(y, v) = f(x, u+v)$, prove that

$$\frac{\partial y}{\partial u} = \frac{\partial}{\partial v}f(y, v)\Big/\frac{\partial}{\partial y}f(y, v),$$

and deduce that $\partial y/\partial u$ is a function of y alone.

15.51. If f, α, β are differentiable functions of x, y, prove that

$$f - \alpha f_\alpha - \beta f_\beta = \begin{vmatrix} \alpha & \beta & f \\ \alpha_x & \beta_x & f_x \\ \alpha_y & \beta_y & f_y \end{vmatrix} \div \frac{\partial(\alpha, \beta)}{\partial(x, y)}.$$

15.52. If $S = \alpha\beta - \frac{1}{2}\gamma^2$, where α, β, and γ are differentiable functions of x, y and x, y are functions of t such that $S = 0$, prove that

$$\frac{\beta\alpha_t - \alpha\beta_t}{\gamma(\gamma - \alpha\gamma_\alpha - \beta\gamma_\beta)} = \frac{x_t}{S_y}\frac{\partial(\alpha, \beta)}{\partial(x, y)}.$$

15.521. If α, β, γ are linear functions of x and y, prove that

$$\begin{vmatrix} \alpha & \beta & \gamma \\ \alpha_x & \beta_x & \gamma_x \\ \alpha_y & \beta_y & \gamma_y \end{vmatrix}$$

is constant.

15.53. If (x, y) lies on the conic $(a, b, c, f, g, h)(x, y, 1)^2 = 0$, prove that

$$\int \frac{dx}{(lx+my+n)(hx+by+f)} = A \log\frac{PT}{PT'},$$

where PT, PT' are the perpendiculars from a point P of the conic onto the tangents at the ends of the chord $lx+my+n = 0$, and A is a certain constant.

15.6. If u, v, w are the three values of t which satisfy

$$t^3 - (2x+z)t^2 + (3x^2 - yz)t - z^4/y = 0,$$

using the substitution $\xi = u+v+w$, $\eta = uv+vw+wu$, $\zeta = uvw$, prove that

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \frac{2(3x+5y)z^4}{y^2(v-w)(w-u)(u-v)}.$$

15.61. Show that under the transformation $u = x^2-y^2$, $v = 2xy$ the equation

$$y^2\frac{\partial^2 H}{\partial x^2}-x^2\frac{\partial^2 H}{\partial y^2} = x\frac{\partial H}{\partial x}-y\frac{\partial H}{\partial y}$$

becomes

$$\left(u\frac{\partial}{\partial v}-v\frac{\partial}{\partial u}\right)\frac{\partial H}{\partial v} = 0.$$

15.62. Prove that, if $f(x,y)$ is twice differentiable, and if $x+y = (u+v)^n$, $x-y = (u-v)^n$, then

$$u^2(x^2-y^2)\left(\frac{\partial^2 f}{\partial x^2}-\frac{\partial^2 f}{\partial y^2}\right) = (u^2-v^2)\left(\frac{\partial^2 f}{\partial u^2}-\frac{\partial^2 f}{\partial v^2}\right).$$

15.621. If w and z are differentiable functions of x and y prove that

$$\left(\frac{\partial z}{\partial x}\right)_w\left(\frac{\partial w}{\partial y}\right)_x+\left(\frac{\partial z}{\partial y}\right)_x\left(\frac{\partial w}{\partial x}\right)_y = \left(\frac{\partial z}{\partial x}\right)_y\left(\frac{\partial w}{\partial y}\right)_x.$$

15.63. If $u = ax+by+c$ and (x,y) lies on the conic $u^2 = 2kxy$ prove that

$$\int\frac{dx}{u(bu-kx)} = \frac{1}{2ck}\log\frac{y}{x}.$$

15.7. If H is a differentiable function of x, y, show that under the transformation $x = u\cos\phi$, $y = u\sin\phi$ we have

$$\frac{\partial^2 H}{\partial x^2}+\frac{\partial^2 H}{\partial y^2} = \frac{\partial^2 H}{\partial u^2}+\frac{1}{u^2}\frac{\partial^2 H}{\partial \phi^2}+\frac{1}{u}\frac{\partial H}{\partial u}.$$

Hence show that if H is a differentiable function of x, y, z, then under the transformation $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ the expression $\frac{\partial^2 H}{\partial x^2}+\frac{\partial^2 H}{\partial y^2}+\frac{\partial^2 H}{\partial z^2}$ becomes

$$\frac{\partial^2 H}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 H}{\partial \theta^2}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 H}{\partial \phi^2}+\frac{2}{r}\frac{\partial H}{\partial r}+\frac{\cot\theta}{r^2}\frac{\partial H}{\partial \theta}.$$

15.71. If z is a function of x, y and

$$p = \frac{\partial z}{\partial x},\quad q = \frac{\partial z}{\partial y},\quad r = \frac{\partial^2 z}{\partial x^2},\quad s = \frac{\partial^2 z}{\partial x\partial y},\quad t = \frac{\partial^2 z}{\partial y^2},$$

and if $Z = px+qy-z$, $X = p$, $Y = q$ and

$$P = \frac{\partial Z}{\partial X},\quad Q = \frac{\partial Z}{\partial Y},\quad R = \frac{\partial^2 Z}{\partial X^2},\quad S = \frac{\partial^2 Z}{\partial X\partial Y},\quad T = \frac{\partial^2 Z}{\partial Y^2},$$

prove that

$$\frac{r}{T} = -\frac{s}{S} = \frac{t}{R} = rt-s^2.$$

15.72. If $2T = ax^2+by^2+cz^2+2fyz+2gzx+2hxy$

where a, b, c, f, g, h are functions of independent variables u and v, and if

$$p = \frac{\partial T}{\partial x},\quad q = \frac{\partial T}{\partial y},\quad r = \frac{\partial T}{\partial z},$$

prove that

$$px+qy+rz = 2T.$$

Show that x, y, z may each be expressed as a function of p, q, r, u, and v and establish the identities

$$\left(\frac{\partial T}{\partial u}\right)_{p,q,r} = -\left(\frac{\partial T}{\partial u}\right)_{x,y,z}, \qquad \left(\frac{\partial T}{\partial v}\right)_{p,q,r} = -\left(\frac{\partial T}{\partial v}\right)_{x,y,z},$$

$$\left(\frac{\partial T}{\partial p}\right)_{q,r} = x, \qquad \left(\frac{\partial T}{\partial q}\right)_{r,p} = y, \qquad \left(\frac{\partial T}{\partial r}\right)_{p,q} = z,$$

provided that the determinant Δ $\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$ is not identically zero.

15.8. If s_n is the sum of the nth powers of the roots of the equation $t^3-\xi t^2+\eta t-\zeta = 0$ prove that

$$\frac{\partial(s_{n+1}, s_{n+2}, s_{n+3})}{\partial(\xi,\eta,\zeta)} = -(n+1)(n+2)(n+3)\zeta^n.$$

15.9. If $l(x) = \int\limits_1^x \frac{1}{t}\,dt$, $x > 0$, prove that there is a functional relation between $l(xy)$ and $l(x)+l(y)$ and hence show that $l(xy) = l(x)+l(y)$ provided $x > 0$, $y > 0$.

15.91. If $f(z)$ is its own derivative and $f(0) = 1$ prove that

$$f(x+y) = f(x)f(y).$$

15.92. If $f(x)$ and $g(x)$ are differentiable functions such that $f'(x) = g(x)$, $g'(x) = -f(x)$, and $f(0) = 0$, $g(0) = 1$, prove that

$$f(x+y) = f(x)g(y)+f(y)g(x) \quad \text{and} \quad g(x+y) = g(x)g(y)-f(x)f(y).$$

15.93. Prove that

$$\int\limits_a^{a^2} \frac{1}{x^2+a^2}\,dx = \frac{1}{a}(\tan^{-1}a-\tfrac{1}{4}\pi)$$

and hence show that

$$\int\limits_a^{a^2} \frac{1}{(x^2+a^2)^2}\,dx = \frac{1}{2a^3}\tan^{-1}a-\frac{\pi}{8a^3}-\frac{(1-a)^2}{4a^3(1+a^2)},$$

and deduce the value of $\int\limits_a^{a^2} \frac{1}{(x^2+a^2)^3}\,dx.$

15.931. If $f(x)$ is continuous in (a,b) show that $f(x)$ is the nth derivative of the function

$$\frac{1}{(n-1)!}\int\limits_a^x (x-t)^{n-1}f(t)\,dt$$

15.932. If $f(x,y)$ is continuous in a rectangle R, and

$$|f(x,Y)-f(x,y)| < M|Y-y|,$$

where M is constant, for any (x,y), (x,Y) in R, and if $\phi_0(x) = b$

$$\phi_{n+1}(x) = b+\int\limits_a^x f(x,\phi_n(x))\,dx$$

where (a, b) is a point in R, prove that we can find a rectangle R^* containing (a, b) and contained in R, such that $\phi_n(x)$ is interval convergent in R^*, and that $y = \lim \phi_n(x)$ is the unique solution of the differential equation

$$\frac{dy}{dx} = f(x, y)$$

in R^* for which $y = b$ when $x = a$.

15.94. A function f(x, y) is said to have an *x-point-derivative* θ at (p, q) if

$$\frac{f(P, q) - f(p, q)}{P - p} \to \theta \quad \text{as} \quad P \to p.$$

Similarly, f(x, y) has a y-point-derivative ϕ at (p, q) if

$$\frac{f(p, Q) - f(p, q)}{Q - q} \to \phi \quad \text{as} \quad Q \to q.$$

If f(x, y) has an x-point-derivative u(x, y) and a y-point-derivative v(x, y) at each point (x, y) of a rectangle R, and if u(x, y), v(x, y) are both continuous in R, prove that f(x, y) is differentiable in R.

15.95. If $f(x, y)$ has point-derivatives $f_x(x, y)$, $f_y(x, y)$, and $f_{xy}(x, y)$ in some region R, and if $f_{xy}(x, y)$ is continuous in R, prove that $f_{yx}(x, y)$ exists and equals $f_{xy}(x, y)$ throughout R.

15.96. If $f(x, 0) = f(0, y) = 0$ for all x, y and if $f(x, y) = \dfrac{x^4}{y}\sin\dfrac{y^2}{x}$ for all non-zero values of x, y, evaluate the point-derivatives f_{xy} and f_{yx} at all points (x, y); prove that f_{xy} and f_{yx} are continuous and verify that $f_{xy} = f_{yx}$ everywhere.

XVI

16. If c is a maximum or minimum value of $f(x, y)$ subject to the condition $g(x, y) = 0$, prove that, in general, the curve $g(x, y) = 0$ touches the curve $f(x, y) = c$.

16.1. Prove that if $xy + yz + zx = 1$ then the maximum value of xyz is $\sqrt{3}/9$.

16.2. Find the maximum and minimum values taken by $2x + y$ as the point (x, y) describes the ellipse $19x^2 + 10xy + 7y^2 = 1$.

16.3. A long rectangular metal sheet, 18 inches wide, is used to make a gutter as follows: the outer edges are bent up, to make equal angles θ with the base, about lines parallel to the longer edges distant x inches from the edges. Find x and θ so that the carrying capacity of the gutter may be a maximum.

16.4. Find the stationary values of $x^2 + y^2 + z^2$ on the surface

$$x^2 + y^2 + z^2 = y(z + x) + 1.$$

16.5. If m and n are integers greater than unity, and b is positive, show that the function $x^m y^n\{(m + n + 1)b - x - y\}$ has a maximum, though not at the origin.

16.6. If $x^2y^2 = (x^2 \sin\alpha - p)(y^2 \sin\alpha - p)$, $p > 0$, show that, for positive values of x and y, the function $xy(x^2 + y^2)$ has a minimum when $x = y$,

provided that α lies inside the interval $\{\frac{1}{6}(12n-7)\pi, \frac{1}{6}(12n+1)\pi\}$ for some value of n.

16.7. Prove that if $x+y+z = 1$, $x^2+y^2+z^2 = 1$ then $x^3+y^3+z^3$ lies between $\frac{5}{9}$ and 1, the first being a true minimum and the second a true maximum value.

16.8. If $V(x,y,z) = x^2+y^2+z^2$, prove that the stationary values of V, subject to $ax^2+by^2+cz^2 = 1$, $lyz+mzx+nxy = 0$ satisfy

$$l^{\frac{2}{3}}(1-aV)^{\frac{1}{3}}+m^{\frac{2}{3}}(1-bV)^{\frac{1}{3}}+n^{\frac{2}{3}}(1-cV)^{\frac{1}{3}} = 0,$$

apart from stationary values at points with *two* coordinates zero.

16.9. If $S(x,y) = ax^2+2hxy+by^2+2gx+2fy+c$, $a \neq 0$, and if the conic $S = 0$ is a rectangular hyperbola, prove that the function $S(x,y)$ has both a maximin and a minimax value at the centre of the conic $S = 0$.

16.91. If a, b, c are constants, and ϕ varies, show that the envelope of the family of circles $(x-c)(x-a\cos\phi)+y(y-b\sin\phi) = 0$ is the curve

$$(x^2+y^2-a^2)\{(x-c)^2+y^2\} = (b^2+c^2-a^2)y^2.$$

16.92. Show that the envelope of a family of parabolas, which has a fixed line for directrix and a variable point on a fixed circle for focus, touches each member of the family at the points in which it is cut by the line joining its focus to the centre of the circle.

16.93. If a family of curves is given by $x = x(\alpha, t)$, $y = y(\alpha, t)$ as α varies, show that the envelope is $x = x(\alpha, t)$, $y = y(\alpha, t)$, where α is a function of t satisfying $\partial(x,y)/\partial(\alpha,t) = 0$.

16.94. Show that the envelope of the family of curves $\phi(x,y,a,b) = 0$, where a, b are connected by the differentiable relation $\lambda(a,b) = 0$, satisfies the equations $\phi = 0$, $\lambda = 0$, $\dfrac{\partial(\phi,\lambda)}{\partial(a,b)} = 0$.

Deduce that the envelope of the normals to the curve $\psi(x,y) = 0$ is the evolute of the curve.

16.95. PM is the perpendicular from the point P, $(\cos\theta, \sin\theta)$, to the y-axis. Find the equation of the rectangular hyperbola which touches the circle $x^2+y^2 = 1$ at P and the y-axis at M, and show that the envelope of the hyperbola, as P describes the circle, consists of the circle itself, the y-axis, and the curve traced by the points in which the hyperbola is cut by the line $x = -2\cos^3\theta$.

16.96. If $A > 0$, $B > 0$, and $n > m > 0$, prove that the function

$$A\{x^{-m}+y^{-m}+(x+y)^{-m}\}-B\{x^{-n}+y^{-n}+(x+y)^{-n}\}$$

has a maximum value on the line $y = x$.

XVII

17. Prove that

$$\int_R e^{ax}\sin by\,dxdy = (4e^{\frac{1}{2}a}/ab)\sin^2\tfrac{1}{2}b\,\mathrm{sh}\,\tfrac{1}{2}a,$$

where R is the square $(0,1)(0,1)$.

17.1. Prove that

$$\int_1^2 dx\int_x^{4-x} f(x,y)\,dy = \int_1^2 dy\int_1^y f(x,y)\,dx + \int_2^3 dy\int_1^{4-y} f(x,y)\,dx,$$

and that the triangular region of integration is of unit area with centroid at $(1\frac{1}{3}, 2)$.

17.2. If R is the triangle with sides along the lines $x = 0, y = 0, x+y = 1$, prove $\int\limits_R (x+y-1)^2\,dxdy = \frac{1}{12}$.

17.3. If $F(x)$ is the integral of a function $f(x)$ such that $F(0) = 0$, prove that $\int\limits_C f(x^2+y^2)\,dxdy = \pi F(1)$, where C is the circle $x^2+y^2 = 1$.

17.4. Find the area bounded by the four parabolas $x^2 = p_1 y$, $x^2 = p_2 y$, $y^2 = q_1 x$, $y^2 = q_2 x$; $0 < p_2 < p_1$; $0 < q_2 < q_1$.

17.5. By integrating $2a/(a^2-y^2\sin^2 x)$ over the rectangle $0 \leqslant x \leqslant \frac{1}{2}\pi$, $0 \leqslant y \leqslant b$, where $b < a$, prove

$$\int\limits^{\frac{1}{2}\pi} \log\frac{a+b\sin x}{a-b\sin x}\,\frac{dx}{\sin x} = \pi \arcsin\frac{b}{a}.$$

17.6. If R is the interior of the ellipse $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2} = 1$ (excluding the foci), and if r_1, r_2 are the distances of any point (x,y) from the foci, show by means of the transformation $x = c\operatorname{ch} u\cos v$, $y = c\operatorname{sh} u\sin v$, where $c^2 = a^2-b^2$, that

$$\int \left(\frac{1}{r_1}+\frac{1}{r_2}\right) dR = 4\pi b.$$

17.7. Show that in the repeated integrals

$$\int\limits^{\infty} dx \int\limits^{\infty} \frac{x-y}{(x+y)^3}\,dy, \qquad \int\limits^{1} dx \int\limits^{1} \frac{x-y}{(x+y)^3}\,dy$$

inverting the order of integration changes the value of the integral.

17.8. If $u_n = \displaystyle\int\limits_0^{\infty} \frac{x^n}{n!}e^{-x}\,dx$, prove that $u_n^2 < u_{2n+1}$.

SOLUTIONS TO EXAMPLES

I

1. If $x \geqslant 0$, $|1+x| = 1+x = 1+|x|$; if $-1 < x < 0$,

$$|1+x| = 1+x = 1-(-x) = 1-|x| < 1+|x|;$$

if $x < -1$, write $-x = 1+y$, $y > 0$, then

$$|1+x| = |-y| = y < 1+(1+y) = 1+|x|.$$

If $x \geqslant 0$, $y \geqslant 0$, $|xy| = xy = |x||y|$; if $x > 0$, $y < 0$ then $xy = -x|y|$ and so $|xy| = x|y| = |x||y|$ etc. Hence $|x/y||y| = |(x/y)y| = |x|$, and so $|x/y| = |x|/|y|$.

It follows that $|1+y/x| \leqslant 1+|y/x|$, $|x|.|1+y/x| \leqslant |x|(1+|y/x|)$ and so $|x+y| \leqslant |x|+|y|$.

Furthermore, $|x| = |(x-y)+y| \leqslant |x-y|+|y|$, whence $|x-y| \geqslant |x|-|y|$, and $|y| = |y-x+x| \leqslant |y-x|+|x|$, whence $|x-y| = |y-x| \geqslant |y|-|x|$, and therefore $|x-y| \geqslant ||x|-|y||$.

1.01. Since $x = +|x|$ or $x = -|x|$, and $x > -|x|$ therefore $x = +|x|$ and so $x > -x$, i.e. $2x > 0$, whence $x > 0$.

1.1. $\left|\sum_1^1 a_r\right| = \sum_1^1 |a_r|$ and by Ex. 1,

$$\Big|\sum_1^{n+1} a_r \qquad \sum_1^{n} a_r + a_{n+1} \qquad \sum_1^{n} a_r\Big| + |a_{n+1}|,$$

and so if $\sum^n a_r \leqslant \sum_1^n |a_r|$ then $\left|\sum_1^{n+1} a_r\right| \leqslant \sum_1^{n+1} |a_r|$.

1.11. $s_n - l = 0(r)$, $n \geqslant \lambda(r)$, and $1/p_n \leqslant 1/\lambda(r)$, $n \geqslant \mu(r)$, so that

$$s_{p_n} - l = 0(r), \qquad n \geqslant \mu(r).$$

1.2. If $a_n^{1/n} \to l < 1$ then $a_n^{1/n} < k < 1$, $l < k$, $n \geqslant n_0$, and so $a_n < k^n$, whence the convergence of $\sum a_n$ follows by comparison with the convergent geometric series $\sum k^n$.

Further $(|x^n|/n^n)^{1/n} = |x|/n \to 0$ for any x, so that $\sum x^n/n^n$ is absolutely convergent.

1.201. $||s_n|-|l|| \leqslant |s_n - l| \to 0$.

1.21. Let $s_n = 1+1/2+1/3+\dots+1/n$, then for any n,

$$0 < s_{2n} - s_n = 1/(n+1)+1/(n+2)+\dots+1/2n > n/2n = 1/2,$$

which proves that s_n is divergent. Furthermore,

$$1/(2^{r-1}+1)+1/(2^{r-1}+2)+\dots+1/2^r > 1/2^r+\dots+1/2^r = 2^{r-1}/2^r = 1/2,$$

and so

$$s_{2^n} = 1+\sum_{r=1}^{n}\{1/(2^{r-1}+1)+1/(2^{r-1}+2)+1/(2^{r-1}+3)+\dots+1/2^r\} > 1+n/2 > N$$

for $n \geqslant 2N-1$.

1.211. $n/(n^2+1) > 1/2n$ so that $\sum n/(n^2+1)$ diverges by comparison with $\sum 1/n$.

Further, if $a_n = 1/n(n+1)$ then $a_{n+1}/a_n = n/(n+2) \to 1$, therefore $\sum x^n/n(n+1)$ is absolutely convergent for $|x| < 1$ and divergent for $|x| > 1$. If $x = 1$ the series becomes $\sum 1/n(n+1)$; but

$$1/n(n+1) = 1/n - 1/(n+1)$$

so that $\sum_1^n 1/r(r+1) = 1-1/(n+1) \to 1$ and $\sum 1/n(n+1)$ is convergent; if $x = -1$ the series becomes $\sum (-1)^n/n(n+1)$ which by the foregoing, is absolutely convergent.

1.22. Let $u_{n+1} = n!/(x+1)(x+2)\ldots(x+n)$, $u_1 = 1$, then $(x+n)u_{n+1} = nu_n$ and so $(x-1)u_{n+1} = nu_n-(n+1)u_{n+1}$, whence

$$(x-1)\sum_1^n u_{r+1} = 1-(n+1)!/(x+1)\ldots(x+n);$$

therefore if $x \neq 1$ then $\sum_1^n u_{r+1} = 1/(x-1)-(n+1)!/(x-1)(x+1)\ldots(x+n)$.

If $y > 0$ then

$$(1+y/2)(1+y/3)\ldots(1+y/n) > 1+y(1/2+1/3+\ldots+1/n) > Ny/2$$

if $n > 2^N$, and so

$$n!/(x+1)(x+2)\ldots(x+n-1) = 1/(1+y/2)(1+y/3)\ldots(1+y/n) \to 0,$$

$y = x-1 > 0$. Thus for $x > 1$ the series converges to $1/(x-1)$. If $x < 1$, write $1-x = z > 0$, then

$$\begin{aligned} n!/(x+1)(x+2)\ldots(x+n-1) &= n!/(2-z)(3-z)\ldots(n-z) \\ &= 1/(1-z/2)(1-z/3)\ldots(1-z/n) \end{aligned}$$

$$> (1+z/p)\{1+z/(p+1)\}\ldots(1+z/n)/(1-z/2)(1-z/3)\ldots\{1-z/(p-1)\}, \quad p > z,$$

$$\text{for} \quad 1/(1-z/r) > 1+z/r, \quad r > z,$$

$$> z\{1/p+1/(p+1)+\ldots+1/n\}/(1-z/2)(1-z/3)\ldots\{1-z/(p-1)\}$$

and so the sequence $n!/(x+1)\ldots(x+n)$ diverges, and therefore $\sum u_r$ diverges.

1.221. If $a_n/b_n \to l > 0$ then $\frac{3}{2}l > a_n/b_n > \frac{1}{2}l$, $n \geq r$, i.e. $\frac{3}{2}lb_n > a_n \geq \frac{1}{2}lb_n$; when $\sum a_n$ converges $a_{n+1}+a_{n+2}+\ldots+a_{n+p} = O(k)$, and so

$$b_{n+1}+b_{n+2}+\ldots+b_{n+p} = (2/l)O(k)$$

so that $\sum b_n$ converges, and similarly, when $\sum b_n$ converges, so does $\sum a_n$. If $\sum a_n$ diverges then we can find $\alpha > 0$ so that for a certain p_n

$$a_{n+1}+a_{n+2}+\ldots+a_{n+p_n} > \alpha,$$

and therefore $b_{n+1}+b_{n+2}+\ldots+b_{n+p_n} > (2/3l)\alpha$, so that $\sum b_n$ diverges. Similarly, the divergence of $\sum a_n$ follows from that of $\sum b_n$.

1.23. $(r-1)/n(n+1)\ldots(n+r-1)$

$$= 1/n(n+1)\ldots(n+r-2)-1/(n+1)(n+2)\ldots(n+r-1),$$

therefore $\sum 1/n(n+1)\ldots(n+r-1)$ converges to $1/(r-1)^2(r-2)!$ Since $n(n+1)\ldots(n+r-1)/n^r \to 1$ the convergence of $\sum 1/n^r$ follows by 1.221.

1.3. Let $s_n = a_0 - a_1 + a_2 - \dots + (-1)^n a_n$, then

$$s_{n+p} - s_n \quad (-1)^{n+1}\{a_{n+1} - a_{n+2} + \dots + (-1)^{p-1} a_{n+p}\};$$

but

$$a_{n+1} - a_{n+2} + a_{n+3} - \dots + (-1)^{p-1} a_{n+p} = (a_{n+1} - a_{n+2}) + (a_{n+3} - a_{n+4}) + \dots +$$

ending with $(a_{n+p-1} - a_{n+p})$ or $+a_{n+p}$ according as p is even or odd. Since a_n decreases, each bracket is positive and so the sum is positive. Furthermore,

$$a_{n+1} - a_{n+2} + a_{n+3} - \dots + (-1)^{p-1} a_{n+p}$$
$$= a_{n+1} - (a_{n+2} - a_{n+3}) - (a_{n+4} - a_{n+5}) - \dots,$$

ending with $-(a_{n+p-1} - a_{n+p})$ or $-a_{n+p}$ according as p is odd or even, which is less than a_{n+1}. Thus $|s_{n+p} - s_n| < a_{n+1} \to 0$, and $\sum (-1)^n a_n$ converges.

1.31. Since $a_n > a_{n+1}$, $n \geqslant N$, therefore a_n steadily decreases, and

$$a_N/a_{N+p} = (a_N/a_{N+1})(a_{N+1}/a_{N+2})\dots(a_{N+p-1}/a_{N+p}) \geqslant (1+k/N)\{1+k/(N+1)\}$$
$$\dots\{1+k/(N+p-1)\} > 1+k\{1/N+1/(N+1)+\dots+1/(N+p-1)\} > \tfrac{1}{2}kn,$$

if $p > N(2^n - 1)$, and so $a_{N+p}/a_N \to 0$, i.e. $a_n \to 0$, since N is fixed.

1.32. Since $\sum a_n$ converges, therefore for $n \geqslant n_k$, and $N > n$

$$a_{n+1} + a_{n+2} + \dots + a_N < 1/10^k;$$

but a_n steadily decreases and so

$$a_{n+1} + a_{n+2} + \dots + a_N > (N-n)a_N = Na_N\left(1 - \frac{n}{N}\right).$$

Hence if $N > 2n$, so that $1 - n/N > \frac{1}{2}$, then $Na_N < 2/10^k$ so that $Na_N \to 0$.

1.4. For $n > 10$, $\frac{1}{2}(n!) > 3.4\ 5.6.7.8.9.10^{n-9} > 10^{n-4}$; but

$$1/(n+1)! + 1/(n+2)! + \dots < (1 + 1/2 + 1/2^2 + \dots)/(n+1)! < 2/(n+1)!$$
$$< 1/10^{n-3} \leqslant 1/10^{k+1} \quad \text{if} \quad n \geqslant k+4.$$

$\sum 1/r! = 2{\cdot}718281828\dots$.

1.401. If $u_n = (n!)^2/(2n)!$ then

$$u_{n+1}/u_n = (n+1)^2/(2n+1)(2n+2) = (1+1/n)^2/4(1+1/n)(1+1/2n) \to \tfrac{1}{4}.$$

1.402. Let $a_n = 1/n^{\frac{1}{2}}$ then $a_n/a_{n+1} = (1+1/n)^{\frac{1}{2}} = p_n$, say. Hence $p_n^2 = 1+1/n$, therefore $p_n - 1 \leqslant 1/k(p_n+1) < 1/2k$, $n \geqslant k$, so that $p_n \to 1$. Thus $\sum a_n t^n$ is absolutely convergent for $|t| < 1$, and divergent for $|t| > 1$. When $t = 1$, $\sum a_n t^n = \sum 1/n^{\frac{1}{2}}$, which diverges since $1/n^{\frac{1}{2}} > 1/n$, and when $t = -1$, $\sum a_n t^n = \sum (-1)^n/n^{\frac{1}{2}}$, which converges since $1/n^{\frac{1}{2}}$ steadily decreases with limit zero. Taking $t = 2x$, we have $\sum 2^n x^n/n^{\frac{1}{2}}$ is absolutely convergent for $|x| < \frac{1}{2}$, convergent for $x = -\frac{1}{2}$, and divergent for $x \geqslant \frac{1}{2}$ or $x < -\frac{1}{2}$.

1.41. $\sum a_{r_n}$ is said to be a rearrangement of $\sum a_n$ if each r_n is an integer, r_n and r_m being different for different n and m, and if to any p corresponds k_p such that $r_{k_p} = p$. Let $s_n = a_0 + a_1 + \dots + a_n$, $t_n = a_{r_0} + a_{r_1} + a_{r_2} + \dots + a_{r_n}$, $m_n = \max(k_0, k_1, k_2, \dots, k_n)$, $q_n = \max(r_0, r_1, r_2, \dots, r_n)$. Then $m_n \geqslant n$, and

for $m \geqslant m_n$, t_m contains all $a_0, a_1, \ldots, a_n$ and perhaps further a_p's for $p > n$, and so $|t_m - s_n| < |a_{n+1}| + |a_{n+2}| + \ldots + |a_{q_m}| \to 0$ since $\sum |a_n|$ converges. Hence if $s_n \to s$ we have $|t_m - s| - |s_n - s| \to 0$ and $|s_n - s| \to 0$ so that $t_m \to s$.

1.5. For instance, $xyz = \lim(x)_n(y)_n(z)_n = \lim(y)_n(z)_n(x)_n = yzx$, etc.

1.501. $(p/q)^m.(p/q)^n = (p^m/q^m)(p^n/q^n) = p^m.p^n/q^m.q^n = p^{m+n}/q^{m+n}$

$$= (p/q)^{m+n}, \text{ etc.}$$

1.51. For any n, $0 < x_n < x_0 + 1$ and so

$$(x_N)^p - (x_n)^p = (x_N - x_n)(x_N^{p-1} + x_N^{p-2}x_n + \ldots + x_n^{p-1})$$
$$< p(x_0+1)^{p-1}/10^n, \quad N \geqslant n,$$

which proves that $(x_n)^p$ converges for a fixed p.

1.511. For fixed m and n, $(x_k)^m \to x^m$, $(x_k)^n \to x^n$ by definition. Hence $(x_k)^m.(x_k)^n \to x^m.x^n$. But, by 1.5, $(x_k)^m\ (x_k)^n = (x_k)^{m+n} \to x^{m+n}$ so that

$$x^m.x^n = x^{m+n}.$$

Further, for $n = 1$, $(x^m)^n = x^m = x^{mn}$, and if $(x^m)^p = x^{mp}$ then $(x^m)^{p+1} = x^m.(x^m)^p = x^m.x^{pm} = x^{m(p+1)}$, whence by induction $(x^m)^n = x^{mn}$ for any n. Similarly, for $m = 1$, $(xy)^m = xy = x^m y^m$, and if $(xy)^p = x^p y^p$ then $(xy)^{p+1} = (xy)(xy)^p = xy.x^p y^p = x^{p+1}y^{p+1}$, so that $(xy)^m = x^m y^m$ for all m.

1.6. $\phi(2^n+1)+\phi(2^n+2)+\ldots+\phi(2^{n+1}) \geqslant 2^n\phi(2^{n+1})$ and

$$\phi(2^n)+\phi(2^n+1)+\ldots+\phi(2^{n+1}-1) \leqslant 2^n\phi(2^n).$$

If $\sum 2^n\phi(2^n)$ converges then $2^n\phi(2^n)+2^{n+1}\phi(2^{n+1})+\ldots \to 0$ and so

$$\phi(2^n)+\phi(2^n+1)+\phi(2^n+2)+\ldots \to 0,$$

which proves $\sum \phi(n)$ converges, if $\sum 2^n\phi(2^n)$ is divergent then we can find α and p_n so that $2^{n+1}\phi(2^{n+1})+2^{n+2}\phi(2^{n+2})+\ldots+2^{p_n}\phi(2^{p_n}) \geqslant \alpha$ and so $\phi(2^n+1)+\phi(2^n+2)+\ldots+\phi(2^{p_n}) \geqslant \tfrac{1}{2}\alpha$, which proves that $\sum \phi(n)$ diverges. The same inequalities prove the converse theorems that $\sum 2^n\phi(2^n)$ converges when $\sum \phi(n)$ converges, and diverges when $\sum \phi(n)$ diverges.

1.61. If $\phi(n) = 1/n^\sigma$, then $\phi(n)$ decreases steadily if $\sigma > 0$.

$$2^n\phi(2^n) = 2^n/2^{\sigma n} = (1/2^{\sigma-1})^n,$$

so that $\sum 2^n\phi(2^n)$ converges if $\sigma > 1$ (for $2^{\sigma-1} > 1$ if $\sigma > 1$) and diverges if $\sigma < 1$. If $\sigma \leqslant 0$, $1/n^\sigma \geqslant 1$ and so $1/n^\sigma$ does not tend to zero.

1.62. Let $\rho = v_m/u_m \leqslant v_{m+1}/u_{m+1} \leqslant v_{m+2}/u_{m+2}\ldots$ so that $v_n \geqslant \rho u_n$, $u_n \leqslant \rho^{-1}v_n$, $n \geqslant m$. If $\sum v_n$ converges then

$$u_n+u_{n+1}+u_{n+2}+\ldots+u_{n+p} \leqslant \rho^{-1}(v_n+v_{n+1}+\ldots+v_{n+p}) \to 0$$

so that $\sum u_n$ converges; and if $\sum u_n$ diverges then we can find α and p_n so that $v_n+v_{n+1}+\ldots+v_{p_n} \geqslant \rho(u_n+u_{n+1}+\ldots+u_{p_n}) \geqslant \rho\alpha$ and therefore $\sum v_n$ diverges.

1.63. If $u_n/u_{n+1} = 1+\beta/n+\theta_n/n^{1+\lambda}$, $|\theta_n| \leqslant M$, $\lambda > 0$, $\beta < 1$, then

$$n(u_n/u_{n+1}-1) \to \beta < 1$$

and $n(u_n/u_{n+1}-1) < 1$, $n \geqslant m$, i.e. $u_n/u_{n+1} \leqslant (1/n)/(1/(n+1))$ whence the divergence of $\sum u_n$ follows from that of $\sum 1/n$.

If $\beta > 1$, then

$$\begin{aligned} nu_n/(n+1)u_{n+1} &= (n/(n+1))+\beta/(n+1)+\theta_n\{n/(n+1)\}/n^{1+\lambda} \\ &= 1+\{(\beta-1)n/(n+1)+\theta_n[n/(n+1)]/n^{\lambda}\}/n = 1+b_n/n, \quad \text{say;} \end{aligned}$$

but $b_n = (\beta-1)n/(n+1)+\{n/(n+1)\}\theta_n/n^{\lambda} \to \beta-1 > 0$, and so for $n \geqslant m$, $nu_n/(n+1)u_{n+1} \geqslant 1+\frac{1}{2}(\beta-1)/n$, whence by Ex. 1.31, $nu_n \to 0$.

Since $nu_n/(n+1)u_{n+1}-1 = b_n/n$, where $b_n \to \beta-1 > 0$, therefore

$$nu_n/u_{n+1}-(n+1) = (1+1/n)b_n > b_n > \tfrac{1}{2}(\beta-1) > 0, \quad n \geqslant m,$$

and so
$$nu_n-(n+1)u_{n+1} > \tfrac{1}{2}(\beta-1)u_{n+1}.$$

Hence

$$\tfrac{1}{2}(\beta-1)(u_{n+1}+u_{n+2}+\dots+u_{n+p}) < nu_n-(n+p)u_{n+p} < nu_n \to 0$$

and therefore $\sum u_n$ is convergent.

1.64.
$$|na_n/(n+1)a_{n+1}| = |a_n/a_{n+1}|/(1+1/n) \to R.$$
$$|(a_n/(n+1))/(a_{n+1}/(n+2))| = |a_n/a_{n+1}|(1+1/(n+1)) \to R.$$

1.7. Since $b_{n+1}-a_{n+1} < k(b_n-a_n)$, therefore

$$b_n-a_n < k^n(b_0-a_0) \to 0, \quad 0 < k < 1.$$

But for any p, (a_{n+p}, b_{n+p}) is contained in (a_n, b_n), therefore

$$0 \leqslant a_{n+p}-a_n \leqslant b_n-a_n \to 0,$$

and similarly $b_{n+p}-b_n \to 0$ so that both a_n and b_n are convergent. If $a_n \to l$ then $b_n-l = (b_n-a_n)+(a_n-l) \to 0$ so that b_n also tends to l.

1.71. $a_{n+1}^2-b_{n+1}^2 = \frac{1}{4}(a_n+b_n)^2-a_nb_n = \frac{1}{4}(a_n-b_n)^2 > 0$ so that $a_n > b_n$ for all n.

$$a_{n+1}-a_n = \tfrac{1}{2}(b_n-a_n) < 0$$

therefore a_n decreases, and

$$b_{n+1}-b_n = \sqrt{b_n}(a_n-b_n)/(\sqrt{a_n}+\sqrt{b_n}) > 0$$

so that b_n increases, and therefore (b_{n+1}, a_{n+1}) is contained in (b_n, a_n) for all n.

Furthermore,

$$a_n-a_{n+1} = (a_n-b_n)-(a_{n+1}-b_{n+1})+(b_n-b_{n+1})$$

and $a_n-a_{n+1} = \frac{1}{2}(a_n-b_n)$ so that $\frac{1}{2}(a_n-b_n)-(a_{n+1}-b_{n+1}) = b_{n+1}-b_n > 0$, i.e. $a_{n+1}-b_{n+1} < \frac{1}{2}(a_n-b_n)$, whence the result follows by 1.7.

1.72. All $a_n > 0$. If $a_0^2-a_0 = k$ then $a_n = a_0$ for all n, and so a_n is equal to the positive root of $x^2-x = k$. If $a_0^2-a_0 < k$ (so that a_0 is less than the positive root of $x^2-x = k$), choose b_0 greater than the positive root of $x^2-x = k$ so that $b_0^2-b_0 > k$, and define b_n by $b_{n+1} = +\sqrt{(b_n+k)}$. Since $a_{n+1}^2 = a_n+k$ therefore $a_{n+2}^2-a_{n+1}^2 = a_{n+1}-a_n$; but

$$a_1^2-a_0^2 = -(a_0^2-a_0-k) > 0$$

and so $a_{n+1} > a_n$ for all n. Similarly $b_{n+1} < b_n$ for all n. But $b_0 > a_0$ and $b_{n+1}^2 - a_{n+1}^2 = b_n - a_n$ so that $b_n > a_n$ for all n, and

$$b_{n+1}-a_{n+1} = (b_n-a_n)/(b_{n+1}+a_{n+1}) < (b_n-a_n)/2a_1, \quad 2a_1 > 2a_0 > 1,$$

so that a_n and b_n tend to the same limit l (say) by 1.7. Since $a_{n+1}^2 = a_n+k$ and $a_{n+1}^2 \to l^2$, $a_n+k \to l+k$ therefore $l^2-l-k = 0$ and $l > 0$. If $a_0^2-a_0 > k$ then the roles of a_n and b_n are interchanged.

1.73. All a_n are positive. If $a_0^2+a_0 = k$ then $a_n = a_0$ for all n.

If $a_0^2+a_0 < k$ then

$$a_1-a_0 = \{k/(1+a_0)\}-a_0 = -(a_0^2+a_0-k)/(1+a_0) > 0 \quad \text{and so} \quad a_1 > a_0.$$

$a_{n+2} = k/(1+a_{n+1}) = k(1+a_n)/(1+k+a_n)$ and so

$$a_{2n+1}-a_{2n} = \{k(1+a_{2n-1})/(1+k+a_{2n-1})\}-\{k(1+a_{2n-2})/(1+k+a_{2n-2})\}$$
$$= k^2(a_{2n-1}-a_{2n-2})/(1+k+a_{2n-2})(1+k+a_{2n-1}),$$

thus $a_{2n+1}-a_{2n}$ has the same sign as $a_{2n-1}-a_{2n-2}$ and is therefore positive since $a_1-a_0 > 0$. Also $a_{2n+1}-a_{2n} < \{k/(1+k)\}^2(a_{2n-1}-a_{2n-2})$.

Furthermore,

$$a_2-a_0 = -(a_0^2+a_0-k)/(1+k+a_0) > 0,$$
$$a_3-a_1 = k(a_0^2+a_0-k)/\{(1+a_0)^2(1+k)+(1+a_0)k\} < 0,$$

and

$$a_{2n+2}-a_{2n} = k^2(a_{2n}-a_{2n-2})/(1+k+a_{2n})(1+k+a_{2n-2}),$$
$$a_{2n+3}-a_{2n+1} = k^2(a_{2n+1}-a_{2n-1})/(1+k+a_{2n+1})(1+k+a_{2n-1}),$$

so that for all n, $a_{2n+2} > a_{2n}$, $a_{2n+3} < a_{2n+1}$, whence

$$a_{2n} < a_{2n+2} < a_{2n+3} < a_{2n+1},$$

i.e. (a_{2n+2}, a_{2n+3}) is contained in (a_{2n}, a_{2n+1}) for all n.

1.74.
$$(a-x_{n+1})/(a+x_{n+1}) = \{(a-x_n)/(a+x_n)\}^2$$
$$= \{(a-x_0)/(a+x_0)\}^{2^{n+1}} \to 0.$$

Similarly
$$(y_{n+1}-a)/(y_{n+1}+a) = \{(y_0-a)/(y_0+a)\}^{2^{n+1}} \to 0.$$

1.75. x_n, $n \geqslant 2$, is uniquely determined when x_0, x_1 are given. If $x_n = t_n$ and $x_n = u_n$ satisfy $x_{n+1} = \frac{1}{2}(x_n+x_{n+1})$ then so does $x_n = at_n+bu_n$ for any constants a, b. For $n \geqslant 2$, the equation is satisfied by $x_n = \alpha^n$ if $\alpha^2 = \frac{1}{2}(\alpha+1)$, i.e. if $\alpha = 1$ or $\alpha = -\frac{1}{2}$ and so it is satisfied by

$$x_n = a+b(-\tfrac{1}{2})^n.$$

Choose a, b so that $x_n = a+b(-\frac{1}{2})^n$ also for $n = 0, 1$; we find

$$a = (x_0+2x_1)/3, \qquad b = 2(x_0-x_1)/3.$$

Hence $x_n \to (x_0+2x_1)/3$.

1.76. Since $|b_{n+2}-b_{n+1}|/|b_{n+1}-b_n| \leqslant k < 1$, therefore $\sum (b_{n+1}-b_n)$ is absolutely convergent, and so $\sum_{0}^{n-1} (b_{r+1}-b_r) = b_n-b_0$ is convergent.

Hence, since $(c_{n+2}-c_{n+1})/(c_{n+1}-c_n) = q/p > 0$, therefore

$$|c_{n+2}-c_{n+1}|/|c_{n+1}-c_n| = q/p < 1$$

so that c_n converges: let λ be the limit of c_n.

From $pc_{n+2}-qc_{n+1} = pc_{n+1}-qc_n$ it follows that

$$pc_{n+2}-qc_{n+1} = pc_1-qc_0$$

and so $$pc_1-qc_0 = \lim(pc_{n+2}-qc_{n+1}) = (p-q)\lambda.$$

1.77. Let y, y^* correspond to x, x^* then

$$y^*-y = \frac{\alpha x^*+\beta}{x^*+\gamma}-\frac{\alpha x+\beta}{x+\gamma} = \frac{(\alpha\gamma-\beta)(x^*-x)}{(x+\gamma)(x^*+\gamma)}. \tag{i}$$

The values of x for which $y = x$ satisfy

$$x^2-(\alpha-\gamma)x-\beta = 0$$

and so λ and μ are self-corresponding; taking $y^* = \lambda$ and $y^* = \mu$ in turn in (i) we find

$$\frac{y-\lambda}{y-\mu} = \frac{\gamma+\mu}{\gamma+\lambda}\cdot\frac{x-\lambda}{x-\mu}.$$

Hence since $a_{n+1} = (pa_n+q^2)/(a_n+p)$ and the roots of the equation

$$x^2-(p-p)x-q^2 = 0$$

are $x = \pm q$, therefore $$\frac{a_{n+1}-q}{a_{n+1}+q} = \frac{p-q}{p+q}\cdot\frac{a_n-q}{a_n+q}$$

and so $$\frac{a_n-q}{a_n+q} = \left(\frac{p-q}{p+q}\right)^n\frac{a_0-q}{a_0+q}.$$

If $p/q > 0$ then

$$\left|\frac{p-q}{p+q}\right|^n = \left|\frac{p/q-1}{p/q+1}\right|^n \to 0, \quad\text{since}\quad \left|\frac{p/q-1}{p/q+1}\right| < 1,$$

and so $a_n \to +q$.

If $p/q < 0$ then

$$\left|\frac{p+q}{p-q}\right|^n = \left|\frac{p/q+1}{p/q-1}\right|^n \to 0, \quad\text{since}\quad \left|\frac{p/q+1}{1-p/q}\right| < 1,$$

and so $a_n \to -q$.

If $p = 0$, then $a_{n+1} = q^2/a_n$, and so $a_{n+2} = a_n$, whence $a_{2n} = a_0$, $a_{2n+1} = a_1 = q^2/a_0$; therefore a_n converges only if $a_0^2 = q^2$, and if $a_0 = q$ then $a_n = q$ for all n, and if $a_0 = -q$ then $a_n = -q$ for all n.

1.8. Let $|p|+|q| = r$ and let $\mu_n = \max(|a_0|, |a_1|,\ldots, |a_n|)$. Then

$$|a_{n+1}| \leqslant r\mu_n.$$

If $r > 1$, then $\mu_{n+1} = \max(\mu_n, |a_{n+1}|) \leqslant r\mu_n$ and so $|a_n| \leqslant \mu_n \leqslant r^n|a_0|$ and therefore $\sum a_n x^n$ converges for $|x| \leqslant 1/R$, $R > r$. If $r \leqslant 1$, then $|a_{n+1}| \leqslant \mu_n \leqslant \mu_0$ and $\sum a_n x^n$ converges for $|x| \leqslant 1/R$, $R > 1$. In the interval of convergence, if $s(x)$ is the limit, then

$$\begin{aligned}(1+px+qx^2)s(x) &= \sum a_n x^n+p\sum a_n x^{n+1}+q\sum a_n x^{n+2}\\ &= a_0+(a_1+pa_0)x+\sum(a_{n+2}+pa_{n+1}+qa_n)x^{n+2} = a_0+(a_1+pa_0)x.\end{aligned}$$

1.9. If $f(x)$ is increasing, for any x and $n > x$ we have $f(n) > f(x)$, whence $l \geqslant f(x)$. Hence $0 \leqslant l-f(n) < 1/k$, $n \geqslant \nu_k$. But $1/g(n) < 1/N$ for $n \geqslant n_N$ and so $g(n) > \nu_k$ for $n \geqslant n_{\nu_k}$, whence $0 \leqslant l-f\{g(n)\} < 1/k$, $n \geqslant n_{\nu_k}$ i.e. $f\{g(n)\} \to l$. Similarly, if $f(x)$ is decreasing.

The example of a function $\phi(x)$ such that $\phi(x) = 0$ if x is a whole number, $\phi(x) = 1$ if $2n < x < 2n+1$, and $\phi(x) = -1$ if $2n+1 < x < 2n+2$, with $g(n) = n+\frac{1}{2}$, illustrates the necessity for the condition '$f(x)$ is increasing (or decreasing)' for $\phi(n) = 0$ and so $\phi(n) \to 0$, but $\phi\{g(2n)\} = 1$ and $\phi\{g(2n+1)\} = -1$ so that $\phi\{g(n)\}$ is not convergent.

II

2. Let $x \geqslant 0$, $X \geqslant 0$, $X-x = 0(2k+1)$; if $x \leqslant 1/10^{2k+1} < 1/10^{2k}$ then $X \leqslant 2/10^{2k+1} < 1/10^{2k}$ and so $\sqrt{x} < 1/10^k$, $\sqrt{X} < 1/10^k$, whence

$$|\sqrt{X}-\sqrt{x}| < 1/10^k;$$

if $x > 1/10^{2k+1}$ then

$$|\sqrt{X}-\sqrt{x}| = |(X-x)/(\sqrt{X}+\sqrt{x})| < |(X-x)/\sqrt{x}| < 10^k/10^{2k+1} = 1/10^{k+1}.$$

2.1. If $q > p^2$, then $x^2+2px+q = (x+p)^2+(q-p^2) > 0$.

2.2. $a_r x^r$ is continuous for any r and the sum of any number of continuous functions is continuous.

2.3. $x = 0$ gives $a_0 = b_0$; then

$$x(a_1+a_2x+\ldots+a_nx^{n-1}) = x(b_1+b_2x+\ldots+b_nx^{n-1})$$

for *all* x near $x = 0$. Hence provided $x \neq 0$

$$a_1+a_2x+\ldots+a_nx^{n-1} = b_1+b_2x+\ldots+b_nx^{n-1}.$$

This equation holds for all values of x near $x = 0$ except perhaps $x = 0$ and so, since its members are continuous, it holds also for $x = 0$, whence $a_1 = b_1$ and so on, and therefore $a_0+a_1x+\ldots+a_nx^n = b_0+b_1x+\ldots+b_nx^n$ for all values of x.

2.4. Divide (a,b) into a finite number of intervals i_k, $k = 1, 2,\ldots, n_\delta$ such that $|f(X)-f(x)| < \frac{1}{2}\delta$ for any x, X in the same interval i_k. Let c_k be the mid-point of i_k; then *either* $f(c_k) > \delta$ *or* $g(c_k) > \delta$. If the former then $f(x) > \frac{1}{2}\delta$ for all x in i_k, and if the latter then $g(x) > \frac{1}{2}\delta$ for all x in i_k

2.5. For any positive λ, and any α, β in $[a,b]$, $\{\lambda g(\alpha)+g(\beta)\}/(\lambda+1)$ lies between $g(\alpha)$ and $g(\beta)$ and therefore, since $g(x)$ is continuous, is a value of $g(x)$ for an x between α and β and so different from zero. Suppose that $g(\alpha) > 0$; take $\lambda = |g(\beta)|/g(\alpha)$, then $g(\beta) \neq -|g(\beta)|$ and so by 1.01, $g(\beta) > 0$. Similarly, if $g(\alpha) < 0$ then $g(\beta) < 0$.

2.6. Let x be any endless decimal in (a,b), so that $a \leqslant x \leqslant b$; but for any n, $(x)_n-1/10^n < x < (x)_n+1/10^n$ and so

$$(x)_n-1/10^n < b, \qquad (x)_n+1/10^n > a.$$

If $a \leqslant (x)_n \leqslant b$ for any n then $f\{(x)_n\} = \lambda$, and since by continuity $f\{(x)_n\} \to f(x)$ therefore $f(x) = \lambda$. If $(x)_n > b$ for some n, then $(x)_n > b$ for all greater values of n, and $(x)_n-1/10^n$ (which is less than b) is greater than a for an n such that $1/10^n < (b-a)$, and so $f\{(x)_n-1/10^n\} = \lambda$; but $f\{(x)_n-1/10^n\} \to f(x)$ and so again $f(x) = \lambda$. Similarly, if $(x)_n < a$, $\lambda = f\{(x)_n+1/10^n\} \to f(x)$, i.e. $f(x) = \lambda$.

If $f(x) = g(x)$ then $f(x)-g(x) = 0$, which being true for terminating decimals is true for all decimals, $f(x)-g(x)$ being continuous.

2.7. We have $a_{n+1}-a_n = -(a_n^2+a_n-k)/(1+a_n)$; if $a_0^2+a_0-k < 0$ then

$$a_{2n+1} > a_{2n}, \qquad a_{2n+1} > a_{2n+2} \quad \text{(see 1.73)}$$

and so $a_{2n}^2+a_{2n}-k$ is negative, $a_{2n+1}^2+a_{2n+1}-k$ is positive, whence it follows since x^2+x-k is continuous that a root of x^2+x-k lies between a_{2n} and a_{2n+1}, etc.

2.8. We have $||f(X)|-|f(x)|| \leqslant |f(X)-f(x)| < 1/10^k$ for any x, X in same sub-interval $(a_r^k, a_{r+1}^k]$, so that $|f(x)|$ is semi-continuous.

Furthermore, for any x in $(a,b]$ we can find r so that $a_r^k \leqslant x < a_{r+1}^k$; hence

$$|f(x)-f(a)| = \left|\sum_{s=0}^{r-1}\{f(a_{s+1}^k)-f(a_s^k)\}+\{f(x)-f(a_r^k)\}\right|$$
$$< \sum_{s=0}^{\nu_k-1}|f(a_{s+1}^k)-f(a_s^k)|+1/10^k = K, \quad \text{say.}$$

Thus $f(x)$ lies between $f(a)-K$ and $f(a)+K$, where K is a constant.

2.81. We can find r, s such that $a_r^k \leqslant c < a_{r+1}^k$ and $a_s^k \leqslant d < a_{s+1}^k$ and therefore $c, a_{r+1}^k, a_{r+2}^k, \ldots, a_s^k, d$ is a k-chain of $f(x)$ in $(c,d]$.

2.82. If $f(x)$ is continuous in (a,b) we can find a_r^k, $r = 0, 1, \ldots, \nu_\kappa$ such that $f(X)-f(x) = 0(k)$ for any x, X in same *closed* sub-interval (a_r^k, a_{r+1}^k) and therefore, *a fortiori*, for any x, X in the same *half-open* sub-interval $(a_r^k, a_{r+1}^k]$.

2.83. Let a_r^k, $r = 0, 1, \ldots, \mu_k$ and b_r^k, $r = 0, 1, 2, \ldots, \nu_k$ be k-chains of the functions $f(x)$ and $g(x)$ respectively. Then if c_r^k, $r = 0, 1, \ldots, \lambda_k$ consists of all the points a_r^k, b_r^k, it follows that c_r^k, $r = 0, 1, \ldots, \lambda_k$ is a k-chain for both $f(x)$ and $g(x)$. Hence if X, x both lie in $(c_r^k, c_{r+1}^k]$, then

$$\{f(X)+g(X)\}-\{f(x)+g(x)\} = \{f(X)-f(x)\}+\{g(X)-g(x)\}$$
$$= 0(k)+0(k) = 0(k-1),$$

and so $f(x)+g(x)$ is semi-continuous; similarly

$$|f(X)g(X)-f(x)g(x)|$$
$$= |f(X)\{g(X)-g(x)\}+g(x)\{f(X)-f(x)\}| < M(0(k)+0(k)) = M0(k-1),$$

where M is a bound of $|f(x)|$ and $|g(x)|$ in $(a,b]$, so that $f(x)g(x)$ is semi-continuous.

III

[illegible] $= \dfrac{\sqrt{(1+x^4)}+x\sqrt{2}}{1-x^2} = \dfrac{1-x^2}{\sqrt{(1+x^4)}-x\sqrt{2}}$,

[illegible] $y-\dfrac{1}{y} = \dfrac{2x\sqrt{2}}{1-x^2} \qquad y+\dfrac{1}{y} = \dfrac{2\sqrt{(1+x^4)}}{1-x^2}$

and therefore $\left(y+\dfrac{1}{y}\right)\dfrac{1}{y}y' = \left(1+\dfrac{1}{y^2}\right)y' = 2\sqrt{2}\left\{\dfrac{1+x^2}{(1-x^2)^2}\right\}$

[illegible]nce

$$D\log y = \frac{1}{y}\frac{dy}{dx} = 2\sqrt{2}\left\{\frac{1+x^2}{(1-x^2)^2}\right\}\frac{(1-x^2)}{2\sqrt{(1+x^4)}} = \frac{(1+x^2)\sqrt{2}}{(1-x^2)\sqrt{(1+x^4)}}.$$

3.1. Since $p(a) = 0$, $p(x)$ is divisible by $x-a$; let $p(x) = (x-a)q(x)$ the[illegible] $p'(x) = (x-a)q'(x)+q(x)$. But $p'(a) = 0$ and so $q(a) = 0$, whence $q(x)$ i[illegible] divisible by $x-a$.

3.2. If $x \geqslant 1$,

$$f\{g(x)\} = f([2x-1]/2x) = 1/2[1-\{1-1/2x\}] = x;$$

if $-1 \leqslant x \leqslant 1$, $\quad f\{g(x)\} = f(\tfrac{1}{2}x) = 2(\tfrac{1}{2}x) = x;$

if $x \leqslant -1$,

$$f\{g(x)\} = f(-[2x+1]/2x) = -1/2[1-(2x+1)/2x] = x.$$

On the other hand, if $\frac{1}{2} \leqslant x < 1$ then $g\{f(x)\} = x$; if $-\frac{1}{2} \leqslant x \leqslant \frac{1}{2}$, $g\{f(x)\} = x$; if $x \geqslant 1$, $g\{f(x)\} = g(0) = 0 \neq x$; if $-1 < x \leqslant -\frac{1}{2}$, $g\{f(x)\} = x$ and if $x \leqslant -1$, $g\{f(x)\} = 0 \neq x$. Observe that $g(x)$ is monotonic increasing but its values are confined to the interval $[-1, +1]$, and so $f\{g(x)\} = x$ for all x implies $g\{f(x)\} = x$ only for x in $[-1, +1]$.

3.3. If p, q, r, s are differentiable functions of x then the derivative of $\begin{vmatrix} p & q \\ r & s \end{vmatrix}$, is $p's-q'r+ps'-qr' = \begin{vmatrix} p' & q' \\ r & s \end{vmatrix} + \begin{vmatrix} p & q \\ r' & s' \end{vmatrix}$. Assuming the theorem for a determinant of the nth order, then if $P_1, P_2, \ldots, P_{n+1}$ are the co-factors of $p_1, p_2, \ldots, p_{n+1}$ in a determinant Δ_{n+1} of the $(n+1)$th order, so that $\Delta_{n+1} = p_1 P_1 + p_2 P_2 + \ldots + p_{n+1} P_{n+1}$, it follows that

$$\Delta'_{n+1} = p'_1 P_1 + p'_2 P_2 + \ldots + p'_{n+1} P_{n+1} + p_1 P'_1 + p_2 P'_2 + \ldots + p_{n+1} P'_{n+1},$$

whence since $P_1, P_2, \ldots, P_{n+1}$ are determinants of the nth order, we deduce the theorem for determinants of the $(n+1)$th order.

3.4. Observe that $u'_r = ru_{r-1}$, $r = 0, 1, 2, 3, 4$ and so the derivative of $u_0^2 u_3 - 3u_0 u_1 u_2 + 2u_1^3$ is $3u_0^2 u_2 - 3u_0^2 u_2 - 6u_0 u_1^2 + 6u_1^2 u_0 = 0$, etc.

3.5. Since $f(x)$ is differentiable $g(x) = f'(a) + 0(r+1)$ for any x such that $0 \leqslant x-a \leqslant 1/10^k$, k depending on r, and so $g(X)-g(x) \equiv 0(r)$ for any x, X in $(a, a+1/10^k)$. Furthermore, $f(x)$ being differentiable, is continuous in (a, b), and in $(a+1/10^k, b)$, $x-a \geqslant 1/10^k$, and therefore $\{f(x)-f(a)\}/(x-a)$ is continuous in $(a+1/10^k, b)$. Hence we can determine λ depending on r so that $g(X)-g(x) = 0(r)$, for any x, X in $(a+1/10^k, b)$ satisfying $|X-x| \leqslant 1/10^\lambda$. Accordingly, $g(X)-g(x) = 0(r)$ for any (x, X) in (a, b) such that $|X-x| \leqslant 1/10^{k+\lambda}$.

3.6. If $\phi(x) = g\{\lambda(x)\}$ then $\phi(x)$ is continuous; but $g(x)$ has a continuous inverse $g^{-1}(y)$, y lying in $\{g(a), g(b)\}$, and so $\lambda(x) = g^{-1}\{\phi(x)\}$ which is continuous in (a, b).

3.7. $(1+x+x^2)^{-1} = (1-x)(1-x^3)^{-1} = (1-x)(1+x^3+x^6+\ldots)$, etc.

3.8. If X and x are both not greater than unity then

$$\{F(X)-F(x)\}/(X-x) = X+x = 2x+(X-x).$$

If X and x are both not less than unity then

$$\{F(X)-F(x)\}/(X-x) = 4-2x-(X-x);$$

if $X > 1 > x$, let $X-1 = \alpha > 0$, $1-x = \beta > 0$, then

$$\begin{aligned}\{F(X)-F(x)\}/(X-x) &= (4X-X^2-2-x^2)/(X-x) \\ &= -\{\alpha^2+\beta^2-2(\alpha+\beta)\}/(\alpha+\beta) = 2-(\alpha+\beta)+2\alpha\beta/(\alpha+\beta).\end{aligned}$$

Hence

$$|\{F(X)-F(x)\}/(X-x)-2x|$$
$$= |2\beta-(\alpha+\beta)+2\alpha\beta/(\alpha+\beta)| < 2(\alpha+\beta) = 2(X-x)$$

and similarly if $X < 1 < x$ and $x-1 = \alpha > 0$, $1-X = \beta > 0$ then

$$|\{F(X)-F(x)\}/(X-x)-(4-2x)|$$
$$= |2\alpha-(\alpha+\beta)+2\alpha\beta/(\alpha+\beta)| < 2(\alpha+\beta) = 2(x-X).$$

Thus if $\phi(x) = 2x$ when $x \leqslant 1$ and $\phi(x) = 4-2x$ when $x > 1$, we have, for any pair x, X

$$|\{F(X)-F(x)\}/(X-x)-\phi(x)| < 2|X-x|,$$

whence $F(x)$ is differentiable in any interval, with derivative $\phi(x)$.

3.81. $f(x)$ is continuous for $x \leqslant 1$, since x is continuous, and $f(x)$ is continuous for $x > 1$, since $2-x$ is continuous. If $x \leqslant 1$ and $X > 1$, then

$$f(X)-f(x) = 2-X-x = (1-x)-(X-1) = 0(p-1)$$

if $1-x = 0(p)$ and $X-1 = 0(p)$. Thus $f(x)$ is continuous in any interval. But

$$\frac{f(X)-f(1)}{X-1} = \frac{(2-X)-1}{X-1} = -1 \quad \text{and} \quad \frac{f(x)-f(1)}{x-1} = 1$$

and so
$$\frac{f(x)-f(1)}{x-1} - \frac{f(X)-f(1)}{X-1} = 2,$$

however close X, x may be. Accordingly $f(x)$ is not differentiable in an interval which contains the point $x = 1$ in its interior. (Note that $f(x)$ is differentiable in an interval (a,b) if $a < b \leqslant 1$ or $1 < a < b$, since x and $2-x$ are differentiable in any interval.)

3.9. The function $F\{f(x)\}$ is differentiable in $\{F(A), F(B)\}$ with derivative $F'\{f(x)\}f'(x) = 1$, and so $F\{f(x)\}-x$ is constant; when $x = 0$, $F\{f(x)\}-x = 0$, and therefore $F\{f(x)\} = x$ for all x in $\{F(A), F(B)\}$.

Furthermore, since $F'(x) > 0$ in (A, B) therefore $F(A) \leqslant F(x) \leqslant F(B)$ in (A, B), whence $F\{f(F(x))\} = F(x)$ in (A, B); since $F(x)$ is monotonic increasing it follows that $f\{F(x)\} = x$ for all x in (A, B), which completes the proof.

3.91. By periodicity,

$$\frac{f(X+a)-f(x+a)}{(X+a)-(x+a)} = \frac{f(X)-f(x)}{X-x}$$

and so $f'(x+a) = f'(x)+0(p)$ for any p, whence result follows.

IV

4.1. (a) $y = \log x/\log y$, therefore $y \log y = \log x$ and $\dfrac{dy}{dx}(\log y+1) = \dfrac{1}{x}$,

whence
$$\frac{dy}{dx} = 1/x(1+\log y).$$

(b) $y\log x = x\log y$, therefore $\dfrac{dy}{dx}\left(\log x - \dfrac{x}{y}\right) = \log y - \dfrac{y}{x}$,

whence $$\frac{dy}{dx} = (yx\log y - y^2)/(xy\log x - x^2).$$

(c) $e^y = \log x/(\log x + \log y)$, $(\log x + \log y)e^y = \log x$, and so

$$\frac{dy}{dx} = y(e^{-y}-1)/x\{1+y(\log x+\log y)\}.$$

4.2. If for some a, $f(a) = 0$ then $f(x) = f(x-a)f(a) = 0$ for all x. Thus either $f(x) = 0$ for all values of x or $|f(x)| > 0$ for all x; in the latter case, differentiating $f(x+y) - f(x)f(y)$ first with respect to x and then with respect to y, we have $f'(x+y) = f'(x)f(y)$ and $f'(x+y) = f(x)f'(y)$, whence $f'(x)/f(x) = f'(y)/f(y) = f'(0)/f(0) = k$ (say) and so $D_x(\log f(x) - kx) = 0$, and therefore $\log f(x) - kx = \log f(0)$; but taking $x = y = 0$ in $f(x+y) = f(x)f(y)$ we have $f(0) = f(0)^2$, whence since $|f(0)| > 0$, $f(0) = 1$, and so $\log f(x) = kx$, i.e. $f(x) = e^{kx}$. Similarly $yg'(xy) = g'(x)$, $xg'(xy) = g'(y)$, whence

$$xg'(x) = yg'(y) = g'(1) = k, \quad \text{say},$$

and so $g(x) - k\log x = g(1)$; but $g(1) = g(1)+g(1)$, so that $g(1) = 0$, whence $g(x) = k\log x$.

4.3. $n^3 = n(n-1)(n-2)+3n(n-1)+n$ and so

$$\sum n^3/n! = \sum 1/n! + 3\sum 1/n! + \sum 1/n! = 5e;$$

similarly

$$n^k = a_1 n + a_2 n(n-1) + a_3 n(n-1)(n-2) + \ldots + a_k n(n-1)\ldots(n-k+1),$$

where $a_k = 1$ and, for any r, $1 < r \leqslant k$, a_r is the remainder when

$$n^{k-1} - a_1 - a_2(n-1) - \ldots - a_{r-1}(n-1)(n-2)\ldots(n-r+2)$$

is divided by $(n-1)(n-2)\ldots(n-r)$, so that each a_r is an integer.

4.4. $f'(x) = -e^{-x}x^n/n! < 0$, $\quad g'(x) = e^{-x}x^n(m-x)/(n-m+1)(n!) > 0$, and so $f(x)$ decreases and $g(x)$ increases; but $f(0) = g(0) = 1$, and so $g(x) > 1 > f(x)$ and in particular $e^m g(m) > e^m > e^m f(m)$. Furthermore e^m differs from $\sum\limits_0^n m^r/r!$ by less than

$$m^{n+1}/(n-m+1)(n!) < m^{n+1}/(n!) < m^{n+1}/10^{n-4}$$

which is less than $1/10^{p+1}$ if $(10/m)^{n+1} > 10^{p+6}$, i.e. if

$$n+1 > (p+6)/(1-\log_{10} m).$$

4.5. $\log(1+1/n)/(1/n) = \log(1+1/n)^n \to \log_e e = 1$. If $\phi(n) = 1/n(\log n)^r$ then $\phi(n)$ steadily decreases and $\sum 2^n\phi(2^n) = (\sum 1/n^r)/(\log 2)^r$.

4.51. Since $D\log(1+t) = \dfrac{1}{1+t}$ therefore $\dfrac{\log(1+t)-\log 1}{t} = 1+O(p)$ when $t = O(q)$; therefore $\dfrac{\log(1+s_n)}{s_n} \to 1$, i.e. $(1+s_n)^{1/s_n} \to e$. Similarly,

$$\log(1-t)/t = -1+O(p), \text{ etc.}$$

4.511. Write $a+b/n = s_n$, $s_n/n \to 0$, then $(1+a/n+b/n^2)^n = (1+s_n/n)^n$; let $e_n = (1+s_n/n)^{n/s_n}$, then $e_n \to e$ and so $s_n \log e_n \to a$, whence $e_n^{s_n} \to e^a$.

4.52. $\log(1+u_r)/u_r = \log(1+u_r)^{1/u_r} \to 1$ since $u_r \to 0$, etc.

4.521. If $x = 0$, $x^m = 0$, if $x > 0$, $x^m = e^{m\log x}$. For $x \geqslant 1/e^k$, $\log x$ is continuous and so $e^{m\log x}$ is continuous and we can choose μ_k so that $|X^m - x^m| < 2/km$ for $x \geqslant e^{-k}$, $X \geqslant e^{-k}$, $|X-x| \leqslant \mu_k$. For $0 \leqslant x < e^{-k}$, $x^m < 1/e^{km} < 1/km$ and so $|X^m - x^m| < 2/km$, $0 \leqslant x \leqslant X < e^{-k}$. Hence if $x \geqslant 0$, $X \geqslant 0$, $|X^m - x^m| < 2/km$, $|X-x| \leqslant \min(\mu_k, e^{-k})$ and so x^m is continuous.

4.53. $a_m/a_{m+p+1} < (1+M/m^{1+\lambda})\{1+M/(m+1)^{1+\lambda}\}...\{1+M/(m+p)^{1+\lambda}\}$, and so $\log a_m/a_{m+p+1} < \sum_{r \geqslant m} \log(1+M/r^{1+\lambda})$; write $u_r = M/r^{1+\lambda}$, then $\sum u_r$ converges and so $\sum \log(1+u_r)$ converges, to s (say). Then $a_m/a_{m+p+1} < e^s$, whence $a_{m+p+1} > a_m e^{-s}$ for any p and so a_n does not tend to zero.

4.6. If $\beta > 0$, $a_n/a_{n+1} = 1+(\beta+\theta_n/n^\lambda)/n \geqslant 1+\frac{1}{2}\beta/n$, since $\beta+\theta_n/n^\lambda \to \beta$, and $\sum a_n$ is convergent by Example 1.31. If $\beta \leqslant 0$,

$$a_n/a_{n+1} \leqslant 1+\theta_n/n^{1+\lambda} < 1+M/n^{1+\lambda}.$$

and so a_n does not tend to zero.

4.7. The cases $\beta > 1$, $\beta < 1$ have been considered in 1.63; there remains the case $\beta = 1$. We have

$$(n\log n)a_n/a_{n+1} - (n+1)\log(n+1) = (1+1/n)n\log n - (n+1)\log(n+1) +$$
$$+ (\theta_n \log n)/n^\lambda = -(n+1)\log(1+1/n) + (\theta_n \log n)/n^\lambda;$$

since $|\theta_n| \leqslant M$ and $(\log n)/n^\lambda \to 0$ therefore $(\theta_n \log n)/n^\lambda \to 0$, also

$$\log(1+1/n) \to 0$$

and $n\log(1+1/n) = \log(1+1/n)^n \to \log e = 1$, hence

$$n\log n(a_n/a_{n+1}) - (n+1)\log(n+1) \to -1,$$

and so, for $n \geqslant m$, $n\log n(a_n/a_{n+1}) < (n+1)\log(n+1)$, whence

$$a_n/a_{n+1} < (1/n\log n)/\{1/(n+1)\log(n+1)\}$$

and the divergence of $\sum a_n$ follows from that of $\sum 1/n\log n$.

4.8. We have $\quad \binom{m}{r}\Big/\binom{m}{r+1} = -\{1+(m+1)/r+\theta_r/r^2\},$

where $\qquad \theta_r = m(m+1)/(1-m/r) \to m(m+1);$

thus for sufficiently great values of r the coefficients are alternately positive and negative. Hence when $x = 1$ the series alternates, and by 4.6 we have convergence for $m+1 > 0$ and divergence for $m+1 \leqslant 0$, and when $x = -1$ the series is (ultimately) positive, and by 4.7 we have convergence for $m+1 > 1$, i.e. $m > 0$, and divergence for $m \leqslant 0$.

When $m > -1$, $\sum \binom{m}{r} x^r$ converges for $x = 1$ and so by Theorem 2.6,

$\sum \binom{m}{r} x^r$ is continuous in $(0, 1)$ and $(1+x)^m$ is also continuous in $(0, 1)$. But for any m, $\sum \binom{m}{r} x^r = (1+x)^m$ in $(0, 1]$; let $x_n = n/(n+1)$, then

$$\sum \binom{m}{r} = \sum \binom{m}{r} \lim x_n^r = \lim \sum \binom{m}{r} x_n^r = \lim(1+x_n)^m$$
$$= (1+\lim x_n)^m = (1+1)^m = 2^m.$$

When $m > 0$, $\sum \binom{m}{r} x^r$ converges for $x = -1$ and so $\sum \binom{m}{r} x^r$ is continuous in $(-1, 0)$. By 4.521, $(1+x)^m$ is continuous in $(-1, 0)$, whence as above $\sum \binom{m}{r} (-1)^r = (1-1)^m = 0$.

4.81. If $y > 0$, $e^y > 1+y$, whence $y-\log(1+y) > 0$ and so $Dx\log(1-1/x) = \log\{(x-1)/x\}+1/(x-1) = 1/(x-1)-\log\{1+1/(x-1)\} > 0$.

4.82. If $b_n = n\log(1+1/n)$ then $a_n = b_{n+1}-b_n$ and so

$$\sum_1^n a_n = b_{n+1}-b_1 \to \log e-\log \tfrac{3}{2} = 1-\log \tfrac{3}{2}.$$

4.9. $D(1-x)e^x = -xe^x < 0$ when $x > 0$, and $D(1+x)e^{-x} = -xe^{-x} < 0$, hence for $x > 0$ $(1-x)e^x < e^0 = 1$ and $(1+x)e^{-x} < e^0 = 1$, and so $1-x < e^{-x} < 1/(1+x)$. Then write x^2 for x.

4.91. $2\,\mathrm{sh}\,\tfrac{1}{2}nx\,\mathrm{sh}\,\tfrac{1}{2}(n+1)x/\mathrm{sh}\,\tfrac{1}{2}x$

$$= (e^{\frac{1}{2}nx}-e^{-\frac{1}{2}nx})\{e^{\frac{1}{2}(n+1)x}-e^{-\frac{1}{2}(n+1)x}\}/(e^{\frac{1}{2}x}-e^{-\frac{1}{2}x})$$
$$= \{e^{(n+1)x}-e^x+e^{-nx}-1\}/(e^x-1)$$
$$= e^x(e^{nx}-1)/(e^x-1)-e^{-x}(e^{-nx}-1)/(e^{-x}-1)$$
$$= \sum_1^n e^{rx}-\sum_1^n e^{-rx} = 2\sum_1^n \mathrm{sh}\,rx, \quad \text{etc.}$$

V

5.01. If $x > 0$, $D|x|^3 = Dx^3 = 3x^2$, and if $x < 0$,

$$D|x|^3 = D(-x^3) = -3x^2.$$

If X, $-x$ are positive and negative respectively then

$$\frac{|X|^3-|-x|^3}{X-(-x)} = \frac{X^3-x^3}{X+x} = X^2-xX+x^2-\frac{2x^3}{X+x};$$

hence if $X = 0(k)$, $x = 0(k)$ then $X^2-xX+x^2 = 0(2k-1)$ and

$$\frac{2x^3}{X+x} < \frac{2x^3}{x} = 2x^2 = 0(2k-1)$$

so that
$$\frac{|X|^3-|-x|^3}{X-(-x)} = 0(2k-2).$$

Hence if $\phi(x) = 3x^2$, $x > 0$, and $\phi(x) = -3x^2$, $x < 0$, and $\phi(0) = 0$, then $\phi(x) = O(2k-1)$ when $x = O(k)$, and $\phi(x)$ is the derivative of $|x|^3$ in any interval. (Note that $\phi(x)$ is *continuous*.)

$D|x| = Dx = 1$ if $x > 0$, $D|x| = D(-x) = -1$ if $x < 0$; if $\phi(x) = 1$, $x > 0$, and $\phi(x) = -1$, $x < 0$, then $\phi(x)$ is not continuous, and so $|x|$ has no derivative, in an interval containing the origin.

$D|\log(1+x)| = D\log(1+x) = 1/(1+x)$ if $x > 0$ and

$$D|\log(1+x)| = D\{-\log(1+x)\} = -\frac{1}{1+x}$$

if $x < 0$; but no derivative in an interval containing the origin since $1/(1+0) = 1$, $-1/(1+0) = -1$.

$D|\sin x| = D\sin x = \cos x$, $\sin x > 0$, and $D|\sin x| = D(-\sin x) = -\cos x$, $\sin x < 0$; no derivative in an interval containing the origin since $\cos 0 = 1$, $-\cos 0 = -1$.

$De^{|x|} = De^x = e^{|x|}$, $x > 0$ and $De^{|x|} = De^{-x} = -e^{-x} = -e^{|x|}$ if $x > 0$. No derivative in an interval containing the origin since $e^0 = 1$, $-e^0 = -1$.

Notice that although for example $|x| = x$ when $x = 0$ it does not follow that $D|x| = 1$ when $x = 0$, for $\dfrac{|X|-|0|}{X-0} = \pm 1$ according as X is positive or negative and so there can be no function $\phi(x)$ such that

$$\frac{|X|-0}{X-0}-\phi(0) = O(k)$$

for *all* X near 0.

5.1. $\sin 3x = \sin(x+2x)$, etc., and $3\cos 3x = D_x \sin 3x$.

5.11. $0 < \sin\frac{1}{4}\pi < \sin\frac{1}{2}\pi = 1$ and $1 = \sin^2\frac{1}{4}\pi+\cos^2\frac{1}{4}\pi = 2\sin^2\frac{1}{4}\pi$

since $$\cos\tfrac{1}{4}\pi = \sin(\tfrac{1}{2}\pi-\tfrac{1}{4}\pi) = \sin\tfrac{1}{4}\pi.$$

$2\sin\frac{1}{6}\pi\cos\frac{1}{6}\pi = \sin\frac{1}{3}\pi = \cos(\frac{1}{2}\pi-\frac{1}{3}\pi) = \cos\frac{1}{6}\pi$; but as $\cos\frac{1}{6}\pi > 0$ therefore

$$2\sin\tfrac{1}{6}\pi = 1.$$

$2\sin\frac{1}{12}\pi\cos\frac{1}{12}\pi = \sin\frac{1}{6}\pi = \frac{1}{2}$ and $\sin^2\frac{1}{12}\pi+\cos^2\frac{1}{12}\pi = 1$, whence

$$(\sin\tfrac{1}{12}\pi+\cos\tfrac{1}{12}\pi)^2 = \tfrac{3}{2}, \qquad (\sin\tfrac{1}{12}\pi-\cos\tfrac{1}{12}\pi)^2 = \tfrac{1}{2};$$

but $\tan\frac{1}{12}\pi < \tan\frac{1}{4}\pi = 1$ so that $\cos\frac{1}{12}\pi-\sin\frac{1}{12}\pi > 0$, whence

$$\sin\tfrac{1}{12}\pi+\cos\tfrac{1}{12}\pi = \sqrt{\tfrac{3}{2}}, \qquad \cos\tfrac{1}{12}\pi-\sin\tfrac{1}{12}\pi = \sqrt{\tfrac{1}{2}}, \text{ etc.}$$

5.12. Use $$\sin(n+\tfrac{1}{2})x-\sin(n-\tfrac{1}{2})x = 2\cos nx\sin\tfrac{1}{2}x$$

and $$\cos(n-\tfrac{1}{2})x-\cos(n+\tfrac{1}{2})x = 2\sin nx\sin\tfrac{1}{2}x.$$

5.2. If the formulae are true for $n = p$ then

$$\begin{aligned}\sin(p+1)x &= \sin px\cos x+\cos px\sin x\\ &= \left\{1+\binom{p}{1}\right\}\cos^p x\sin x-\left\{\binom{p}{2}+\binom{p}{3}\right\}\cos^{p-2}x\sin^3 x+\dots\\ &= \binom{p+1}{1}\cos^p x\sin x-\binom{p+1}{3}\cos^{p-2}x\sin^3 x+\dots\end{aligned}$$

and

$$\cos(p+1)x = \cos px\cos x - \sin px \sin x$$
$$= \cos^{p+1}x - \left\{\binom{p}{1}+\binom{p}{2}\right\}\cos^{p-1}x\sin^2 x + \left\{\binom{p}{3}+\binom{p}{4}\right\}\cos^{p-3}x\sin^4 x - \dots$$
$$= \cos^{p+1}x - \binom{p+1}{2}\cos^{p-1}x\sin^2 x + \binom{p+1}{4}\cos^{p-3}x\sin^4 x - \dots$$

whence the result follows by induction.

Observe that both $\cos nx$ and $\sin(n+1)x/\sin x$ contain only even powers of $\sin x$.

5.201. $C_0(\cos\theta) = 1 = \cos 0$, $C_1(\cos\theta) = \cos\theta$; if for $n = 1, 2, 3,\dots, k$, $C_n(\cos\theta) = \cos n\theta$, then

$$C_{k+1}(\cos\theta) = 2\cos\theta C_k(\cos\theta) - C_{k-1}(\cos\theta)$$
$$= 2\cos\theta\cos k\theta - \cos(k-1)\theta = \cos(k+1)\theta.$$

Hence $C_n(\cos\theta) = \cos n\theta$ for all values of n. Similarly

$$S_n(\cos\theta) = \sin(n+1)\theta/\sin\theta$$

for all n.

5.202. If $D_n(x) = S_n(x) - C_n(x)$ then $D_{n+2}(x) = 2xD_{n+1}(x) - D_n(x)$; furthermore, $xS_{n+1}(x) = 2x.xS_n(x) - xS_{n-1}(x)$. Now $D_1(x) = S_1(x) - C_1(x) = xS_0(x)$ and $D_2(x) = S_2(x) - C_2(x) = 2x^2 = xS_1(x)$, and if $D_n(x) = xS_{n-1}(x)$ and $D_{n+1}(x) = xS_n(x)$ then $D_{n+2}(x) = xS_{n+1}(x)$, so that, by induction,

$$D_n(x) = xS_{n-1}(x)$$

for all n. Accordingly $C_n(x) = S_n(x) - xS_{n-1}(x) = xS_{n-1}(x) - S_{n-2}(x)$ and therefore $x^nC_n(x) = x^{n+1}S_{n-1}(x) - x^nS_{n-2}(x)$, whence by addition, since $xC_1(x) = x^2S_0(x)$, $\sum_1^n x^nC_n(x) = x^{n+1}S_{n-1}(x)$. Taking $x = \cos\theta$, we find

$$\sum_1^n \cos^n\theta\cos n\theta = \cos^n\theta\sin n\theta\cot\theta.$$

5.21. Since $\cos n\alpha = \cos n\left(\alpha+\frac{2r\pi}{n}\right)$, $r = 0, 1, 2,\dots, (n-1)$, therefore the equation $\cos nx = \cos n\alpha$ has the solutions $x = \alpha+\frac{2r\pi}{n} = \alpha_r$ (say), $r = 0, 1,\dots, (n-1)$, i.e. the equation $C_n(\cos x) = \cos n\alpha$ has these solutions, whence the solutions of $C_n(y) = \cos n\alpha$ are $y = \cos\left(\alpha+\frac{2r\pi}{n}\right) = \cos\alpha_r$, $r = 0, 1,\dots, n-1$, and so

$$C_n(y) - \cos n\alpha = A(y-\cos\alpha)(y-\cos\alpha_1)(y-\cos\alpha_2)\dots(y-\cos\alpha_{n-1})$$

or

$$\cos nx - \cos n\alpha$$
$$= A(\cos x - \cos\alpha)(\cos x-\cos\alpha_1)(\cos x - \cos\alpha_2)\dots(\cos x - \cos\alpha_{n-1}).$$

To determine the constant A we equate the coefficients of $\cos^n x$ on both sides; if λ_n is the coefficient of x^n in the polynomial of the nth degree

$C_n(x)$ then $\lambda_1 = 1$ and, from 5.201, $\lambda_{n+2} = 2\lambda_{n+1}$ whence $\lambda_n = 2^{n-1}$ and so $A = 2^{n-1}$.

Since $C_n(y) - \cos n\alpha = 2^{n-1}(y-\cos\alpha)(y-\cos\alpha_1)...(y-\cos\alpha_{n-1})$ for all values of y in $(-1, 1)$, by 2.3 the two polynomials are equal for all values of y.

Let $y = \frac{1}{2}\left(x+\frac{1}{x}\right)$ then $C_1(y) = \frac{1}{2}\left(x+\frac{1}{x}\right)$; if $C_n(y) = \frac{1}{2}\left(x^n+\frac{1}{x^n}\right)$, $n = 1, 2, 3,..., k$, then

$$\begin{aligned} C_{k+1}(y) &= 2yC_k(y) - C_{k-1}(y) \\ &= \frac{1}{2}\left(x+\frac{1}{x}\right)\left(x^k+\frac{1}{x^k}\right) - \frac{1}{2}\left(x^{k-1}+\frac{1}{x^{k-1}}\right) = \frac{1}{2}\left(x^{k+1}+\frac{1}{x^{k+1}}\right), \end{aligned}$$

whence $C_n(y) = \frac{1}{2}\left(x^n+\frac{1}{x^n}\right)$ for all n. Accordingly

$$\begin{aligned} &\frac{1}{2}\left(x^n+\frac{1}{x^n}\right) - \cos n\alpha \\ &\qquad = 2^{n-1}\left\{\frac{1}{2}\left(x+\frac{1}{x}\right)-\cos\alpha\right\}\left\{\frac{1}{2}\left(x+\frac{1}{x}\right)-\cos\alpha_1\right\}...\left\{\frac{1}{2}\left(x+\frac{1}{x}\right)-\cos\alpha_{n-1}\right\}, \end{aligned}$$

i.e.

$$\begin{aligned} &x^{2n} - 2x^n\cos n\alpha + 1 \\ &\qquad = (x^2-2x\cos\alpha+1)(x^2-2x\cos\alpha_1+1)(x^2-2x\cos\alpha_2+1)... \\ &\qquad\qquad (x^2-2x\cos\alpha_{n-1}+1). \end{aligned}$$

Writing $\alpha = 2\pi/n$, so that $\alpha_r = 2(r+1)\pi/n$ and $n\alpha = 2\pi$, we obtain

$$\begin{aligned} (x^n-1)^2 = (x-1)^2\left(x^2-2x\cos\frac{2\pi}{n}+1\right)\left(x^2-2x\cos\frac{4\pi}{n}+1\right)... \\ \left(x^2-2x\cos 2(n-1)\frac{\pi}{n}+1\right); \end{aligned}$$

but $\cos 2(n-r)\frac{\pi}{n} = \cos\frac{2r\pi}{n}$ and so

$$x^n - 1 = (x-1)\left(x^2-2x\cos\frac{2\pi}{n}+1\right)...\left(x^2-2x\cos\frac{2m\pi}{n}+1\right)$$

if $n = 2m+1$ and

$$x^n-1 = (x-1)(x+1)\left(x^2-2x\cos\frac{2\pi}{n}+1\right)...\left(x^2-2x\cos 2(m-1)\frac{\pi}{n}+1\right)$$

if $n = 2m$. Taking $\alpha = \pi/n$ we obtain similar expressions for x^n+1.

5.22. If $|x| \leqslant 1$, let $x = \cos\theta$, then

$$D_x\cos n\theta = D_\theta\cos n\theta/D_\theta\cos\theta = n\sin n\theta/\sin\theta$$

and so

$$\begin{aligned} D_x\{x^nC_n(x)\} &= nx^{n-1}\cos n\theta + x^n n\sin n\theta/\sin\theta \\ &= nx^{n-1}(\cos n\theta\sin\theta+\sin n\theta\cos\theta)/\sin\theta \\ &= nx^{n-1}\sin(n+1)\theta/\sin\theta = nx^{n-1}S_n(x). \end{aligned}$$

Since the *polynomials* $D_x\{x^nC_n(x)\}$ and $nx^{n-1}S_n(x)$ are equal for all x in $(-1, 1)$, they are equal everywhere.

5.3. Since

$$\sin^4 x = (\sin^2 x)^2 = \tfrac{1}{4}(1-\cos 2x)^2 = \tfrac{1}{4}\{1-2\cos 2x+\tfrac{1}{2}(1+\cos 4x)\}$$
$$= \tfrac{3}{8}-\tfrac{1}{2}\cos 2x+\tfrac{1}{8}\cos 4x$$

and $\cos 2x$ lies between $1-2x^2+2x^4/3$ and $1-2x^2+2x^4/3-4x^6/45$ and $\cos 4x$ lies between $1-8x^2+32x^4/3$ and $1-8x^2+32x^4/3-256x^6/45$, therefore $\sin^4 x$ lies between

$$\tfrac{3}{8}-(\tfrac{1}{2}-x^2+x^4/3)+(\tfrac{1}{8}-x^2+4x^4/3-32x^6/45) = x^4-32x^6/45$$

and $$\tfrac{3}{8}-(\tfrac{1}{2}-x^2+x^4/3-2x^6/45)+(\tfrac{1}{8}-x^2+4x^4/3) = x^4+2x^6/45.$$

5.301. Since $1 \geqslant \cos\dfrac{x}{n} \geqslant 1-\dfrac{x^2}{2n^2}$, therefore

$$1 \geqslant \cos^n\frac{x}{n} \geqslant \left(1-\frac{x^2}{2n^2}\right)^n \to e^0 = 1,$$

by Example 4.511.

5.31. $$(\sin x/x)^2 > (1-x^2/6)^2 = 1-x^2/3+x^4/36, \text{ if } x^2 < 6.$$
$$\cos x < 1-x^2/2+x^4/24.$$

If $0 < |x| \leqslant \frac{1}{2}\pi$ then $x^2 < 12$ and so $x^2 > x^4/12$ whence

$$x^2(\tfrac{1}{2}-\tfrac{1}{3}) > (x^4/12)(\tfrac{1}{2}-\tfrac{1}{3})$$

and therefore $$1-x^2/3+x^4/36 > 1-x^2/2+x^4/24.$$

5.32. $1/n > \sin 1/n > 1/n-1/6n^3$ and so $n < 1/\sin(1/n) < 6n^3/(6n^2-1)$, whence $0 < \{1/\sin(1/n)\}-n < n/(6n^2-1) \to 0$.

5.33. (i) If $f(x) = \sin x - x\cos x$ then $f(0) = 0$, $f'(x) = x\sin x > 0$ in $[0, \frac{1}{2}\pi)$ so that $f(x)$ *increases* from zero;

(ii) if $f(x) = e^{\frac{1}{2}x^2}\cos x$ then $f(0) = 1$ and $f'(x) = e^{\frac{1}{2}x^2}(x\cos x-\sin x) < 0$, whence $f(x) < 1$, $0 < x \leqslant \frac{1}{2}\pi$.

5.34. $D\{(1-\cos t)/t\} = (t\sin t+\cos t-1)/t^2$, and

$$D(t\sin t+\cos t-1) = t\cos t > 0, \qquad 0 < t < \tfrac{1}{2}\pi,$$

and so $(1-\cos t)/t$ steadily increases as t increases between 0 and $\frac{1}{2}\pi$; writing $t = x/n$, it follows that $n(1-\cos x/n)$ steadily decreases as n increases ($n > 2x/\pi$). Hence since $\lim n\left(1-\cos\dfrac{x}{n}\right) \leqslant \lim\dfrac{x^2}{2n} = 0$, therefore $\sum(-1)^n n(1-\cos x/n)$ converges.

5.4. Take $x = 0$, then $f(0)+f(y) = f(y)$, i.e. $f(0) = 0$. Also

$$f'(x) = f'\{(x+y)/(1-xy)\}(1+y^2)/(1-xy)^2, \quad xy \neq 1,$$

whence $f'(0) = f'(y)(1+y^2)$, i.e. $D_x\{f(x)-f'(0)\tan^{-1}x\} = 0$ and so

$$f(x)-f'(0)\tan^{-1}x = 0.$$

5.6. Since $\cos(n+2)\theta+\cos n\theta = 2\cos\theta\cos(n+1)\theta$, therefore by Example 1.8,

$$1+r\cos\theta+r^2\cos 2\theta+\ldots = \frac{1-r\cos\theta}{1-2r\cos\theta+r^2}$$

and since $\sin(n+2)\theta+\sin n\theta = 2\cos\theta\sin(n+1)\theta$, by Example 1.8,

$$r\sin\theta+r^2\sin 2\theta+\ldots = \frac{r\sin\theta}{1-2r\cos\theta+r^2},$$

both series being absolutely convergent for $|r| < 1$.

5.7. Since $|\cos nx| \leqslant \frac{1}{n^2}$ and $\sum \frac{1}{n^2}$ is convergent, therefore $\sum \frac{\cos nx}{n^2}$ is absolutely convergent. If $|X-x| < 1/N^2$ then

$$\left|\sum \frac{\cos nX}{n^2} - \sum \frac{\cos nx}{n^2}\right|$$

$$\sum \frac{1}{n^2}(\cos nx-\cos nX)+\sum_{N+1} \frac{\cos nx}{n^2}-\sum_{N+1} \frac{\cos nX}{n^2}$$

$$\sum_{1} \frac{2}{n^2}\sin n\frac{(X-x)}{2}\ \ldots\ (X+x)\ +2\sum_{N+1} \frac{1}{n^2}$$

$$\sum_{1}^{N} \frac{2}{n^2}\qquad (X-x)\ +2\sum_{N+1}^{\infty}\left(\frac{1}{n-1}-\frac{1}{n}\right)$$

$$< \sum_{1}^{N} \frac{|X-x|}{n}+\frac{2}{N} < N|X-x|+\frac{2}{N} < \frac{3}{N},$$

so that $\sum \frac{\cos nx}{n^2}$ is continuous in any interval.

VI

6. Differentiating with respect to a the identity

$$\frac{2a}{x^2-a^2} = \frac{1}{x-a}-\frac{1}{x+a},$$

we find
$$\frac{2}{x^2-a^2}+\frac{4a^2}{(x^2-a^2)^2} = \frac{1}{(x-a)^2}+\frac{1}{(x+a)^2},$$

whence, taking $a = 1$,

$$\frac{4}{(x^2-1)^2} = \frac{1}{(x+1)^2}+\frac{1}{x+1}-\frac{1}{x-1}+\frac{1}{(x-1)^2}.$$

VII

7.01. If $2y = x\left(1+\dfrac{dy}{dx}\right)$ then $\dfrac{d^3y}{dx^3} = 0$, $x \neq 0$.

7.02. Since $f'g' = 1$ therefore $f''g'^2+f'^2g'' = 0$, and so
$$h^2 = f^3g+3f^2g'+3f'g^2+g^3f = f^3g+g^3f,$$
whence dividing by h the result follows.

7.1. Differentiate the relation $y(1+x^2) = 2$ three times in succession.

7.11. Denote the derivative of y by y'; then we have
$$ky' = (1+y')\cos(x+y), \qquad ky'' = -(1+y')^2\sin(x+y)+y''\cos(x+y)$$
whence $\qquad ky''(1+y') = -ky(1+y')^3+ky'y''$, etc.
When $x = 0$, $\sin y = ky$; since for $0 < y \leqslant \frac{1}{2}\pi$, $1 > \sin y/y \geqslant 2/\pi$, therefore $\sin y - ky = 0$ has no solution, except $y = 0$, when $k > 1$ or $k < 2/\pi$, and so the condition $k > 1$ ensures that $y = 0$ when $x = 0$. By Theorem 7.4 the coefficients of x, $x^2/2!$, $x^3/3!$, $x^4/4!$ in the expansion of y are the values of dy/dx, d^2y/dx^2, d^3y/dx^3, d^4y/dx^4 when $x = 0$.

7.2. (i) Partial fractions, (ii) Leibnitz's theorem,
$$\text{(iii)} \quad \sin^3x\cos x = \tfrac{1}{4}\sin 2x - \tfrac{1}{8}\sin 4x, \text{ etc.}$$

7.21. Denote d^nz/dx^n by z_n and dP_n/dy by P'_n. Then $z_1 = x\cos y$, $z_2 = \cos y - x^2\sin y = \cos y - 2y\sin y$, and if $z_{2n} = P_{2n}\sin y + Q_{2n}\cos y$ then $z_{2n+1} = x[(P'_{2n}-Q_{2n})\sin y + (Q'_{2n}+P_{2n})\cos y]$, and if
$$z_{2n+1} = x(P_{2n+1}\sin y + Q_{2n+1}\cos y)$$
then
$$\begin{aligned} z_{2n+2} &= \sin y(P_{2n+1}+x^2P'_{2n+1}-x^2Q_{2n+1})+\cos y(Q_{2n+1}+x^2Q'_{2n+1}+x^2P_{2n+1}) \\ &= (P_{2n+1}+2yP'_{2n+1}-2yQ_{2n+1})\sin y+(Q_{2n+1}+2yQ'_{2n+1}+2yP_{2n+1})\cos y \end{aligned}$$
whence the result follows by induction.

Hence from $z_{2n} = P_{2n}\sin y + Q_{2n}\cos y$ we derive, using the foregoing results,
$$\begin{aligned} z_{2n+2} = \sin y\{P'_{2n}-Q_{2n}+2y(P''_{2n}-Q'_{2n}-Q'_{2n}-P_{2n})\}+ \\ +\cos y\{Q'_{2n}+P_{2n}+2y(Q''_{2n}+P'_{2n}+P'_{2n}-Q_{2n})\} \end{aligned}$$
and so $\qquad P_{2n+2} = P'_{2n}-Q_{2n}+2y(P''_{2n}-2Q'_{2n}-P_{2n})$.

7.211. The first derivative of $\tan x$ is $1+\tan^2x$; if for some n the nth derivative of $\tan x$ is $p(\tan x)$, where $p(t)$ is a polynomial of the $(n+1)$th degree, then the $(n+1)$th derivative of $\tan x$ is $p'(\tan x)(1+\tan^2x)$ which is a polynomial of the $(n+2)$th degree, whence the result follows by induction.

7.22. $y_0 = y$,
$$y_1 = e^{x\cos\alpha}\{\cos\alpha\sin(x\sin\alpha)+\sin\alpha\cos(x\sin\alpha)\} = e^{x\cos\alpha}\sin(\alpha+x\sin\alpha),$$
$$y_2 = e^{x\cos\alpha}\sin(2\alpha+x\sin\alpha)$$

and so

$$y_2 - 2y_1 \cos\alpha + y_0$$
$$= e^{x\cos\alpha}\{\sin(2\alpha + x\sin\alpha) + \sin(x\sin\alpha) - 2\cos\alpha\sin(\alpha + x\sin\alpha)\}$$
$$= 2e^{x\cos\alpha}\{\sin(\alpha + x\sin\alpha)\cos\alpha - \sin(\alpha + x\sin\alpha)\cos\alpha\} = 0.$$

Differentiating $y_2 - 2y_1\cos\alpha + y_0 = 0$ n times we have

$$y_{n+2} - 2y_{n+1}\cos\alpha + y_n = 0.$$

When $x = 0$, $y_0 = 0$, $y_1 = \sin\alpha$, $y_2 = \sin 2\alpha$ and if $y_n = \sin n\alpha$ then $y_{n+1} = 2\cos\alpha\sin n\alpha - \sin(n-1)\alpha = \sin(n+1)\alpha$, and so $y_n = \sin n\alpha$ for all n, whence by Theorem 7.4,

$$y = \sum \frac{x^n}{n!}\sin n\alpha.$$

7.23. Use Example 3.1 and induction.

7.3. By 2.5 $f'(x)$ is of constant sign in $[a, b]$; if $f'(x)$ is positive in $[a, b]$ then $f(a)$ is not a maximum and $f(b)$ is not a minimum, i.e. $f(a), f(b)$ are not both maximum nor both minimum. Similarly if $f'(x)$ is negative in $[a, b]$.

7.31. If α, β lie in $[a, b]$ and $f(\alpha), f(\beta)$ are both maximum values of $f(x)$, let $\alpha_1, \alpha_2, \ldots, \alpha_p$ be the points between α and β where $f'(x) = 0$; since $f(\alpha)$ is a maximum, by 7.3, $f'(x)$ is negative in $[\alpha, \alpha_1]$, and similarly $f'(x)$ is positive in $[\alpha_p, \beta]$. Furthermore $f'(x)$ is of constant sign in each interval (α_r, α_{r+1}). Let ρ be the least value of r for which $f'(x)$ is positive in (α_r, α_{r+1}); then $f'(x)$ is negative in $(\alpha_{\rho-1}, \alpha_\rho)$ and positive in $(\alpha_\rho, \alpha_{\rho+1})$, and therefore $f(x)$ has a minimum value at $x = \alpha_\rho$.

7.41. If $y = (a-b)x/(x+a)(x+b) = a/(x+a) - b/(x+b)$, $dy/dx = 0$ when $x = \pm\sqrt{(ab)}$, $d^2y/dx^2 = -(\sqrt{a}-\sqrt{b})/\sqrt{(ab)}(\sqrt{a}+\sqrt{b})^3$ if $x = +\sqrt{(ab)}$, and $d^2y/dx^2 = (\sqrt{a}+\sqrt{b})/\sqrt{(ab)}(\sqrt{a}-\sqrt{b})^3$ if $x = -\sqrt{(ab)}$; hence the maximum value is $(\sqrt{a}-\sqrt{b})/(\sqrt{a}+\sqrt{b})$ and the minimum is $(\sqrt{a}+\sqrt{b})/(\sqrt{a}-\sqrt{b})$.

7.411. If $f(x) = x^{\frac{2}{3}}(x^2-4)$, then $f'(x) = \frac{8}{3}(x^{\frac{5}{3}} - x^{-\frac{1}{3}})$, $f''(x) = \frac{8}{9}(5x^{\frac{2}{3}} + x^{-\frac{4}{3}})$ When $f'(x) = 0$ then $x^2 = 1$, i.e. $x = \pm 1$, and $f''(x)$ is *positive* for any value of x. Thus $f(x)$ is minimum when $x = 1$ and when $x = -1$, and since $f(0) = 0$ and $f(x) < 0$ on either side of the point $x = 0$, $f(0)$ is a maximum value.

7.44. $f(x) = \sin^3 x\cos 3x$,

$$f'(x) = 3\sin^2 x\cos x\cos 3x - 3\sin^3 x\sin 3x = 3\sin^2 x\cos 4x,$$
$$f''(x) = 6\sin x\cos x\cos 4x - 12\sin^2 x\sin 4x$$
$$= 3\sin 2x\cos 4x + 6\cos 2x\sin 4x - 6\sin 4x,$$
$$f'''(x) = 30\cos 2x\cos 4x - 24\sin 2x\sin 4x - 24\cos 4x.$$

When $f'(x) = 0$, $\sin x = 0$ or $\cos 4x = 0$, i.e. $x = n\pi$ or $x = \frac{1}{8}(2m+1)\pi$. When $x = n\pi$, $f''(x) = 0$, $f'''(x) = 6$, so that $x = n\pi$ is a point of inflexion, for all n. When $x = \frac{1}{8}(2m+1)\pi$,

$$f''(x) = 3\sin\tfrac{1}{4}(2m+1)\pi\cos\tfrac{1}{2}(2m+1)\pi - 12\sin^2\tfrac{1}{8}(2m+1)\pi\sin\tfrac{1}{2}(2m+1)\pi,$$

i.e. $$f''(x) = -12\sin^2\tfrac{1}{2}(2m+1)\pi\sin\tfrac{1}{2}(2m+1)\pi;$$
but $\sin\frac{1}{2}(2m+1)\pi$ is $+1$ or -1 according as m is even or odd; therefore $f''(x) < 0$ if m is even and $f''(x) > 0$ if m is odd. Hence $f(x)$ has an inflexion at each point $x = n\pi$, a maximum at each point $\frac{1}{2}(4m+1)\pi$ and a minimum at each point $\frac{1}{2}(4m+3)\pi$.

7.5. Let $AB = x$, $\angle ABC = \theta$, then $x^2\cos\theta$ is constant and $(AC+3BC)^2$ is proportional to $(\sin\theta+3\cos\theta)^2/\cos\theta = \sec\theta+6\sin\theta+8\cos\theta$; the first derivative
$$\sec\theta\tan\theta+6\cos\theta-8\sin\theta = (t^3-7t+6)/(1+t^2)^{\frac{1}{2}}$$
$$= (t+3)(t-1)(t-2)/(1+t^2)^{\frac{1}{2}},$$
where $t = \tan\theta$. When $0 < t < 1$ the derivative is positive, it vanishes at $t = 1$, is negative for $1 < t < 2$, vanishes again at $t = 2$, and is positive for $t > 2$. Hence $(AC+3BC)^2$ is maximum when $t = 1$ and minimum when $t = 2$. Observe that $f(t)^2$ and $f(t)$ are maximum and minimum together, when $f(t) > 0$, for $f(T)^2-f(t)^2 = \{f(T)-f(t)\}\{f(T)+f(t)\}$ and so $f(T)^2-f(t)^2$ and $f(T)-f(t)$ have the same sign.

7.6. Denote the cofactors in Δ of $t_r, u_r, v_r, \ldots$ by $T_r, U_r, V_r, \ldots$ and the cofactors in Γ by $\overline{T}_r, \overline{U}_r, \overline{V}_r, \ldots$. Since a determinant with two identical rows is zero, we have $D_x T_0 = \overline{T}_n$, $D_x U_0 = \overline{U}_n, \ldots$ and $T_0 = \overline{T}_{n+1}$, $U_0 = \overline{U}_{n+1}, \ldots$. But, as $\Gamma = 0$, $\overline{T}_n\overline{U}_{n+1}-\overline{T}_{n+1}\overline{U}_n = 0$, i.e. $(D_x U_0)T_0-(D_x T_0)U_0 = 0$, whence $D_x(U_0/T_0) = 0$, etc.

7.7. $y = \tan^{-1}x$ then $(1+x^2)\dfrac{dy}{dx} = 1$, and differentiate n times by Leibnitz's theorem. From $(1+x^2)y_{n+1}+2nxy_n+n(n-1)y_{n-1} = 0$, taking $x = 0$ and denoting the value of y_n at $x = 0$ by y_n^0 we find $y_{n+1}^0 = -n(n-1)y_{n-1}^0$; but $y^0 = 0$, $y_1^0 = 1$ and so $y_{2n}^0 = 0$, $y_{2n+1}^0 = (-1)^n(2n)!$, whence by Theorem 7.4, $y = \sum(-1)^n x^{2n+1}/(2n+1)$.

7.71. $y = \sin^{-1}x$, $\dfrac{dy}{dx} = 1/\sqrt{(1-x^2)}$, $\dfrac{d^2y}{dx^2} = x/(1-x^2)^{\frac{3}{2}}$ whence
$$(1-x^2)\frac{d^2y}{dx^2} = x\frac{dy}{dx};$$
differentiate n times by Leibnitz's theorem and proceed as in 7.7.

7.8. By the binomial theorem $1/(1+x^2) = 1-x^2+x^4-x^6+\ldots$, $|x| < 1$. Furthermore $\sum(-1)^n x^{2n+1}/(2n+1)$ converges for $-1 < x \leqslant 1$ and so $D_x\sum(-1)^n x^{2n+1}/(2n+1) = \sum(-1)^n x^{2n}$, $|x| < 1$, and $D_x\tan^{-1}x = 1/(1+x^2)$ and so $\tan^{-1}x-\sum(-1)^n x^{2n+1}/(2n+1) = \text{constant} = 0$, $|x| < 1$, and by continuity the result extends to $x = 1$. Similarly
$$D_x\sin^{-1}x = (1-x^2)^{-\frac{1}{2}} = \sum\frac{1.3.5\ldots2n-1}{2.4.6\ldots2n}x^{2n}$$
$$= D_x\sum\frac{1.3.5\ldots2n-1}{2.4.6\ldots2n}\frac{x^{2n+1}}{2n+1}, \text{ etc., } |x| < 1;$$

since $\sum \frac{1.3.5...2n-1}{2.4.6...2n}\frac{x^{2n+1}}{2n+1}$ converges by the Gauss test at $x = \pm 1$, the result extends also to $x = \pm 1$.

7.9. $(px+q)^2/(ax^2+2bx+c)$ is positive or negative for all x near $-q/p$, according as $aq^2-2bpq+cp^2$ is positive or negative.

7.91. $Ax^2+2Bx+C-\lambda(ax^2+2bx+c)$ is a square if

$$\phi(\lambda) = (B-\lambda b)^2-(A-\lambda a)(C-\lambda c) = 0;$$

for large values of λ, $\phi(\lambda)$ and $\phi(-\lambda)$ have the same sign as b^2-ac, which is negative, and $\phi(A/a) = (B-\lambda b)^2 > 0$, so that $\phi(\lambda)$ has a root greater than A/a and a root less than A/a; let λ_1, λ_2 be the roots so chosen that $A-\lambda_1 a > 0$, $A-\lambda_2 a < 0$. Then

$$Ax^2+2Bx+C-\lambda_1(ax^2+2bx+c) = (px+q)^2$$

and $$Ax^2+2Bx+C-\lambda_2(ax^2+2bx+c) = -(rx+s)^2,$$

where $p = \sqrt{(A-\lambda_1 a)}$, $r = \sqrt{(a\lambda_2-A)}$, etc. Hence

$$\{(Ax^2+2Bx+C)/(ax^2+2bx+c)\} = \lambda_1+(px+q)^2/(ax^2+2bx+c) \quad \text{and}$$
$$= \lambda_2-(rx+s)^2/(ax^2+2bx+c),$$

whence by 7.9 result follows.

7.92. Since $P(x)$ is a polynomial of the nth degree, therefore $P(a+h)$ is a polynomial in h of the nth degree, say

$$P(a+h) = c_0+c_1 h+c_2\frac{h^2}{2!}+c_3\frac{h^3}{3!}+...+c_n\frac{h^n}{n!}.$$

Then, differentiating with respect to h,

$$P^k(a+h) = c_k+c_{k+1}h+c_{k+2}\frac{h^2}{2!}+...+c_n\frac{h^{n-k}}{(n-k)!},$$

whence $P^k(a) = c_k$.

7.921. Since $f(x) = g(x)$ for $|x| < r$, and since the power series are differentiable for $|x| < r$, therefore $f^n(x) = g^n(x)$ for $|x| < r$; in particular $f^n(0) = g^n(0)$. But $f^n(0) = a_n$, $g^n(0) = b_n$ and so $a_n = b_n$ for all n. Hence $f(x) = g(x)$ for all x such that $|x| < R$.

VIII

8.01. $1/\{\sqrt{(x+2)}+\sqrt{(x+1)}\} = \sqrt{(x+2)}-\sqrt{(x+1)}$.

8.02. $\int dx/x(1+x^5) = \int x^4\, dx/x^5(1+x^5) = \frac{1}{5}\int dt/t(1+t)$, $\quad t = x^5$.

8.03. $3\sin x+4\cos x = \frac{7}{2}(\sin x+\cos x)+\frac{1}{2}(\cos x-\sin x)$.

8.04. Write $2e^x = y^2+1$, then integral becomes $\int 2dy/(y^2+1)$.

8.05. Write $20x-15 = 49\sin^2\theta$, $64-20x = 49\cos^2\theta$ and integral becomes

$$(49/10)\int\{\sin^2\theta\cos\theta/(5-4\cos\theta)\}\,d\theta = (49/10)\int\{(\cos^2\theta-\cos\theta)/(4\cos\theta-5)\}\,d\theta$$
$$= (49/40)\int\cos^2\theta\,d\theta+(49/32)\int\cos\theta\,d\theta+(441/160)\int d\theta-$$
$$-(441/128)\int d\theta/(5-4\cos\theta).$$

8.06. Write $x = 1-t$ and expand $(1-t)^5$.

8.07. Write
$$y = \frac{\sqrt{f}+\sqrt{g}}{\sqrt{(f-g)}} = \frac{\sqrt{(f-g)}}{\sqrt{f}-\sqrt{g}}$$
so that
$$y+\frac{1}{y} = \frac{2\sqrt{f}}{\sqrt{(f-g)}}, \qquad y-\frac{1}{y} = \frac{2\sqrt{g}}{\sqrt{(f-g)}} = 2(f/g-1)^{-\frac{1}{2}}$$
and therefore
$$\left(y+\frac{1}{y}\right)\frac{1}{y}\frac{dy}{dx} = \left(1+\frac{1}{y^2}\right)\frac{dy}{dx} = -(f/g-1)^{-\frac{3}{2}}(f'g-fg')/g^2$$
$$= (fg'-f'g)/\sqrt{\{g(f-g)^3\}},$$
whence
$$\frac{d}{dx}\log y = \frac{1}{y}\frac{dy}{dx} = \frac{fg'-f'g}{\sqrt{\{g(f-g)^3\}}}\,\frac{\sqrt{(f-g)}}{2\sqrt{f}} = \frac{1}{2\sqrt{(fg)}}\,\frac{fg'-f'g}{f-g}.$$

8.08. Since
$$u_n = \int \frac{1}{(1-x^4)^n}\,dx = \int \frac{1-x^4}{(1-x^4)^{n+1}}\,dx = u_{n+1}-\frac{1}{4}\int x\,\frac{4x^3}{(1-x^4)^{n+1}}\,dx$$
$$= u_{n+1}-\frac{1}{4n}\left\{\frac{x}{(1-x^4)^n}-\int\frac{1}{(1-x^4)^n}\,dx\right\}$$
we have
$$4nu_{n+1} = (4n-1)u_n+x(1-x^4)^{-n}, \quad n \geqslant 1.$$
Since
$$u_1 = \int\frac{1}{1-x^4}\,dx = \frac{1}{4}\int\frac{1}{1-x}\,dx+\frac{1}{4}\int\frac{1}{1+x}\,dx+\frac{1}{2}\int\frac{1}{1+x^2}\,dx$$
$$= \frac{1}{4}\log\left|\frac{x+1}{x-1}\right|+\frac{1}{2}\arctan x,$$
therefore u_n is determined for all n
$$\int\frac{x^4}{(1-x^4)^3}\,dx = u_3-u_2, \quad \text{etc},$$
$$\int\frac{x^8}{(1-x^4)^4}\,dx = -\int\frac{x^4(1-x^4)}{(1-x^4)^4}\,dx+\int\frac{x^4}{(1-x^4)^4}\,dx = -u_3+u_2-u_3+u_4$$
$$= u_2-2u_3+u_4, \quad \text{etc.}$$

8.1. $\int (1+x^2)^{1/4}x^{-15/2}\,dx = \int(1+1/x^2)^{1/4}x^{-4}\,dx/x^3 = -2\int(t^4-1)^2t^4\,dt$, etc.

8.2. $\int(\log x)^{k+1}\,dx = x(\log x)^{k+1}-(k+1)\int(\log x)^k\,dx$.

8.21. We have $g\{g^*(x)\} = x$ and $g'\{g^*(x)\} = 1/D_xg^*(x)$, and therefore, when $t = g^*(x)$,
$$\int f\{g(t)\}g'(t)\,dt = \int f[g\{g^*(x)\}]g'\{g^*(x)\}\{D_xg^*(x)\}\,dx = \int f(x)\,dx.$$

8.3. If
$$y^2 = (x^2-8x+10)/(3x^2-10x+9) = \tfrac{1}{3}[1-(14x-21)/(3x^2-10x+9)]$$
then
$$y\frac{dy}{dx} = 7(x-1)(x-2)/(3x^2-10x+9)^2;$$
hence the maximum and minimum values of y^2 are $y^2 = \frac{3}{2}$ when $x = 1$

and $y^2 = -2$ when $x = 2$. By Example 7.91, $\frac{3}{2}-y^2$ is proportional to $(x-1)^2$ and y^2+2 is proportional to $(x-2)^2$; in fact

$$\tfrac{3}{2}-y^2 = 7(x-1)^2/2(3x^2-10x+9) \quad\text{and}\quad y^2+2 = 7(x-2)^2/(3x^2-10x+9).$$

Now

$$\{(x-4)/(3x^2-10x+9)\sqrt{(x^2-8x+10)}\}\frac{dx}{dy}$$
$$= (x-4)\sqrt{(3x^2-10x+9)}/7(x-1)(x-2)$$
$$= \tfrac{3}{7}\sqrt{(3x^2-10x+9)}/(x-1)-\tfrac{2}{7}\sqrt{(3x^2-10x+9)}/(x-2),$$

writing $(x-4)/(x-1)(x-2)$ in partial fractions,

$$= \tfrac{3}{7}/\sqrt{\{\tfrac{2}{7}[\tfrac{3}{2}-y^2]\}}-\tfrac{2}{7}/\sqrt{\{\tfrac{1}{7}(y^2+2)\}}$$

whence

$$\sqrt{14}\int (x-4)\,dx/(3x^2-10x+9)\sqrt{(x^2-8x+10)}$$
$$= 3\int dy/\sqrt{(\tfrac{3}{2}-y^2)}-2\sqrt{2}\int dy/\sqrt{(y^2+2)}$$
$$= 3\sin^{-1}y\sqrt{(\tfrac{2}{3})}-2\sqrt{2}\log\{y+\sqrt{(y^2+2)}\}.$$

8.31. If $y^2 = (5x^2+2x-7)/(5x^2+12x+8) = 1-5(2x+3)/(5x^2+12x+8)$ then $y\dfrac{dy}{dx} = 25(x+1)(x+2)/(5x^2+12x+8)^2$ and so the maximum value of y^2 is $\frac{9}{4}$, when $x = -2$ and the minimum is -4 when $x = -1$. Then $\frac{9}{4}-y^2 = 25(x+2)^2/4(5x^2+12x+8)$ and $y^2+4 = 25(x+1)^2/(5x^2+12x+8)$.

Hence

$$\{1/(5x^2+12x+8)\sqrt{(5x^2+2x-7)}\}\,dx/dy = \sqrt{(5x^2+12x+8)}/\{25(x+1)(x+2)\}$$
$$= 1/5\sqrt{(y^2+4)}-1/10\sqrt{(\tfrac{9}{4}-y^2)},$$

whence the value of the integral is $\frac{1}{5}\log\{y+\sqrt{(y^2+4)}\}-\frac{1}{10}\sin^{-1}(2y/3)$.

8.4. If $x+\sqrt{(x^2-a^2)} = t$ then

$$x-\sqrt{(x^2-a^2)} = a^2/t \quad\text{and}\quad \log\{x+\sqrt{(x^2-a^2)}\} = \log t,$$
$$\frac{1}{\sqrt{(x^2-a^2)}}\frac{dx}{dt} = \frac{1}{t},$$

whence

$$\int\frac{x+\sqrt{(x^2-a^2)}}{x-\sqrt{(x^2-a^2)}}\frac{dx}{\sqrt{(x^2-a^2)}} = \int\frac{\{x+\sqrt{(x^2-a^2)}\}^2}{a^2}\frac{dx}{\sqrt{(x^2-a^2)}}$$
$$= \frac{1}{a^2}\int t^2\frac{dt}{t} = \frac{t^2}{2a} = \frac{\{x^2+\sqrt{(x^2-a^2)}\}^2}{2a^2}$$

and

$$\int\left(1+\frac{x}{\sqrt{(x^2-a^2)}}\right)^{-3}dx = \int\frac{\sqrt{(x^2-a^2)^4}}{\{x+\sqrt{(x^2-a^2)}\}^3}\frac{dx}{\sqrt{(x^2-a^2)}}$$
$$= \frac{1}{16}\int\left(t-\frac{a^2}{t}\right)^4\frac{1}{t^4}\,dt$$
$$= \frac{1}{16}\int\left(1-\frac{4a^2}{t^2}+\frac{6a^4}{t^4}-\frac{4a^6}{t^6}+\frac{a^8}{t^8}\right)dt$$
$$= \frac{1}{16}\left\{t+\frac{4a^2}{t}-\frac{2a^4}{t^3}+\frac{4}{5}\frac{a^6}{t^5}-\frac{1}{7}\frac{a^8}{t^7}\right\}, \quad\text{etc}$$

IX

9.01. $\int_{1/3}^{3/2} dx/(1+x^2) = \tan^{-1}\frac{3}{2} - \tan^{-1}\frac{1}{3} = \tan^{-1}\frac{7}{9} = 0{\cdot}661.$

$$\int \{x^2/\sqrt{(1-x^2)}\}\,dx = \int \sin^2\theta\,d\theta, \quad x = \sin\theta, \quad 0 < \theta < \tfrac{1}{2}\pi$$

$$= \tfrac{1}{2}\int (1-\cos 2\theta)\,d\theta = \tfrac{1}{2}(\theta - \sin\theta\cos\theta).$$

If $\sin\theta_1 = \frac{3}{5}$, $\sin\theta_2 = \frac{4}{5}$ then $\sin(\theta_2-\theta_1) = \frac{7}{25}$ and so $\theta_2-\theta_1 = 0{\cdot}284$: furthermore, $\sin\theta_1\cos\theta_1 = \sin\theta_2\cos\theta_2$ and so

$$\int_{3/5}^{4/5} \{x^2/\sqrt{(1-x^2)}\}\,dx = \tfrac{1}{2}[\theta - \sin\theta\cos\theta]_{\theta_1}^{\theta_2} = 0{\cdot}142.$$

9.02. $\displaystyle\int \frac{1}{\sqrt{\{x(1-x)\}}}\,dx = \int \frac{2\sin t\cos t}{\sin t\cos t}\,dt = 2t, \quad x = \sin^2 t, \quad 0 < t < \tfrac{1}{2}\pi.$

Therefore, if $0 < a < 1$,

$$\int_{1/n}^{a^2} \frac{1}{\sqrt{\{x(1-x)\}}}\,dx = 2\sin^{-1}a - 2\sin^{-1}(1/\sqrt{n}) \to 2\sin^{-1}a,$$

and

$$\int_{a^2}^{1-(1/n)} \frac{1}{\sqrt{\{x(1-x)\}}}\,dx = 2\sin^{-1}\left\{\sqrt{\left(1-\frac{1}{n}\right)}\right\} - 2\sin^{-1}a \to \pi - 2\sin^{-1}a.$$

Hence $$\int_0^1 \frac{1}{\sqrt{\{x(1-x)\}}}\,dx = \pi - 2\sin^{-1}a + 2\sin^{-1}a = \pi.$$

9.1. $$n\int \sin^{n-1}\theta\sin(n+1)\theta\,d\theta = \sin^n\theta\sin n\theta$$

and $$n\int \{\sin(n-1)\theta/\sin^{n+1}\theta\}\,d\theta = -\sin n\theta/\sin^n\theta.$$

9.101. Since

$$\sin(-7\pi/6) = -\sin(\pi+\tfrac{1}{6}\pi) = \tfrac{1}{2} \quad\text{and}\quad \sin(7\pi/3) = \sin(\tfrac{1}{3}\pi) = \sqrt{\tfrac{3}{4}},$$

and since $\sin t$ is differentiable everywhere, and in particular in $(-7\pi/6, 7\pi/3)$, therefore all the conditions sufficient for the transformation are satisfied.

9.11. $$\int_0^N x^n e^{-x^2}\,dx = \int_0^N \tfrac{1}{2}x^{n-1}2xe^{-x^2}\,dx$$

$$= [-\tfrac{1}{2}x^{n-1}e^{-x^2}]_0^N + \tfrac{1}{2}(n-1)\int_0^N x^{n-2}e^{-x^2}\,dx;$$

since

$$\int_M^N x^n e^{-x^2}\,dx < \int_M^N \{(n+1)!/x^{n+2}\}\,dx = n!(1/M^{n+1} - 1/N^{n+1}) < n!/M^{n+1} \to 0$$

as M increases, therefore $\int_0^\infty x^n e^{-x^2}\,dx$ exists, and furthermore, $N^{n-1}e^{-N^2} \to 0$ as N increases.

9.12. $\int_0^\infty dx/(1+e^x)^p(1+e^{-x})$

$$= \int_0^\infty e^x\,dx/(1+e^x)^{p+1} = [-1/p(1+e^x)^p]_0^\infty = 1/2^p p.$$

$$\int_0^{2a} x\,dx/\surd(2ax-x^2) = \int_0^{2a} a\,dx/\surd\{a^2-(a-x)^2\} - \int_0^{2a} (a-x)\,dx/\surd\{a^2-(a-x)^2\}$$

$$= [a\sin^{-1}(x/a-1)]_0^{2a} - [\surd\{a^2-(a-x)^2\}]_0^{2a} = \pi a.$$

$$\int_0^1 (1-x^2)^n\,dx = \int_0^{\frac{1}{2}\pi} \cos^{2n+1}\theta\,d\theta, \quad x = \sin\theta$$

$$= 2n!!/(2n+1)!!.$$

$$\int^N (1+x^2)^{-n}\,dx = \int^{\tan^{-1}N} \cos^{2n-2}\theta\,d\theta, \quad x = \tan\theta$$

$$\to \int_0^{\frac{1}{2}\pi} \cos^{2n-2}\theta\,d\theta = \{(2n-3)!!/(2n-2)!!\}(\tfrac{1}{2}\pi).$$

9.13. $$1+2(\cos 2x+\cos 4x+\ldots+\cos 2nx) = \sin(2n+1)x/\sin x,$$

whence

$$\int_0^{\frac{1}{2}\pi} \{\sin(2n+1)x/\sin x\}\,dx = \int_0^{\frac{1}{2}\pi} 1\,dx = \tfrac{1}{2}\pi, \quad \text{since} \quad \int_0^{\frac{1}{2}\pi} \cos 2rx\,dx = 0, \quad r > 0.$$

Also $$\sin t+\sin 3t+\ldots+\sin(2n-1)t = (\sin nt)^2/\sin t$$

and so $$\int_0^{\frac{1}{2}\pi} (\sin nt/\sin t)^2\,dt = \sum_{r=1}^{n} \int_0^{\frac{1}{2}\pi} \{\sin(2r-1)t/\sin t\}\,dt = n\pi/2.$$

9.2. $$\int_{-1}^{1} \sin\alpha\,dx/(1-2x\cos\alpha+x^2) = \int_{-1}^{1} \sin\alpha\,dx/\{(x-\cos\alpha)^2+\sin^2\alpha\}$$

$$= [\tan^{-1}\{(x-\cos\alpha)/\sin\alpha\}]_{-1}^{1}$$

$$= \tan^{-1}\{(1-\cos\alpha)/\sin\alpha\}+\tan^{-1}\{(1+\cos\alpha)/\sin\alpha\}$$

$$= \tan^{-1}(\tan\tfrac{1}{2}\alpha)+\tan^{-1}(\cot\tfrac{1}{2}\alpha);$$

if $n\pi < \frac{1}{2}\alpha < n\pi+\frac{1}{2}\pi$, then $\tan^{-1}(\tan\frac{1}{2}\alpha) = \tan^{-1}\{\tan(\frac{1}{2}\alpha-n\pi)\} = \frac{1}{2}\alpha-n\pi$ and $\tan^{-1}(\cot\frac{1}{2}\alpha) = \tan^{-1}(\tan\{\frac{1}{2}\pi-\frac{1}{2}\alpha+n\pi\}) = \frac{1}{2}\pi-\frac{1}{2}\alpha+n\pi$ and value of integral is $\frac{1}{2}\pi$; if $n\pi-\frac{1}{2}\pi < \frac{1}{2}\alpha < n\pi$, $\tan^{-1}(\tan\frac{1}{2}\alpha) = \frac{1}{2}\alpha-n\pi$ and

$$\tan^{-1}(\cot\tfrac{1}{2}\alpha) = \tan^{-1}[\tan\{\tfrac{1}{2}\pi-\tfrac{1}{2}\alpha+(n-1)\pi\}] = \tfrac{1}{2}\pi-\tfrac{1}{2}\alpha+(n-1)\pi$$

and the value of the integral is $-\frac{1}{2}\pi$.

9.201. $\displaystyle\int \frac{1}{(1+x)\sqrt{(1-x)(1-kx)}}\,dx = -\int \frac{du}{\sqrt{(2u-1)\{(k+1)u-k\}}},$

$$u = 1/(x+1),$$

$$= -\sqrt{\frac{2}{(k+1)}}\int dv, \qquad 2u-1 = \frac{(1-k)}{1+k}\operatorname{sh}^2 v, \quad (k+1)u-k = \frac{(1-k)}{2}\operatorname{ch}^2 v,$$

$$= -v\sqrt{\frac{2}{k+1}}.$$

But $\displaystyle\operatorname{ch} v = \sqrt{\frac{2}{1-k}}\{(k+1)u-k\} = \sqrt{\left(\frac{1-kx}{1+x}\right)}\Big/\left(\frac{1-k}{2}\right)$

and $\displaystyle\operatorname{sh} v = \sqrt{\left(\frac{1+k}{1-k}\right)}(2u-1) = \sqrt{\left(\frac{1+k}{1-k}\right)\left(\frac{1-x}{1+x}\right)}$

and so when $x = 1$, $\operatorname{sh} v = 0$, and therefore $v = 0$, and when $x = 0$,

$$v = \log\left\{\sqrt{\left(\frac{2}{1-k}\right)}+\sqrt{\left(\frac{1+k}{1-k}\right)}\right\},$$

whence

$$\int_0^1 \frac{1}{(1+x)\sqrt{(1-x)(1-kx)}}\,dx = \left(\sqrt{\frac{2}{k+1}}\right)\log\left\{\sqrt{\left(\frac{2}{1-k}\right)}+\sqrt{\left(\frac{1+k}{1-k}\right)}\right\}.$$

Taking $k = \cos 4\alpha$ we find

$$\int_0^1 \frac{1}{(1+x)\sqrt{(1-x)(1-x\cos 4\alpha)}}\,dx = \frac{1}{\cos 2\alpha}\log\left(\frac{1}{\sin 2\alpha}+\frac{\cos 2\alpha}{\sin 2\alpha}\right)$$
$$= \sec 2\alpha \log \cot \alpha$$

9.21. $\displaystyle\int dx/x^5\sqrt{(1-x^2)} = \int x\,dx/x^6\sqrt{(1-x^2)} = -\int dy/(1-y^2)^3,$

$$y^2 = 1-x^2,$$

$$= -\tfrac{1}{8}\int dy\{1/(1-y)+1/(1-y)^2+1/(1+y)+1/(1+y)^2\}, \quad \text{etc.}$$

9.3. $\displaystyle I = \int_0^{2l} xf(x)\,dx = \int_0^{2l}(2l-y)f(y)\,dy, \quad y = 2l-x,$

$$\text{since } f(y) = f\{l-(l-y)\} = f\{l+(l-y)\} = f(2l-y),$$

$$= 2l\int_0^{2l} f(y)\,dy - I,$$

whence

$$I = l\int_0^{2l} f(y)\,dy = l\left\{\int_0^l + \int_l^{2l} f(y)\,dy\right\} = lI_1+lI_2, \quad \text{say;}$$

in I_2 write $y = 2l-z$ then $\displaystyle I_2 = -\int_l^0 f(z)\,dz = I_1$ and so $\displaystyle I = 2l\int_0^l f(x)\,dx.$

9.31. $\int \sec x \sec(a-x)\,dx =$ [illegible]

$$= \int_{-a}^{a} dy/(\cos a+\cos y), \quad y = 2x-a,$$

$$= \int_{-\tan\frac{1}{2}a}^{\tan\frac{1}{2}a} dt/(\cos^2\tfrac{1}{2}a - t^2\sin^2\tfrac{1}{2}a), \quad t = \tan\tfrac{1}{2}y,$$

$$= [\operatorname{cosec} a \log\{(1+t\tan\tfrac{1}{2}a)/(1-t\tan\tfrac{1}{2}a)\}]_{-\tan\frac{1}{2}a}^{\tan\frac{1}{2}a}$$

$$= 2\operatorname{cosec} a \log\sec a,$$

or $\sec x\sec(a-x) = \sec^2 x(\cos a+\sin a\tan x)^{-1}$

$$= D\operatorname{cosec} a\log(\cos a+\sin a\tan x).$$

Writing $x = \frac{1}{2}a+y$, $\sec x\sec(a-x) = \sec(\frac{1}{2}a+y)\sec(\frac{1}{2}a-y)$, which is unchanged by replacing y by $-y$, so that $\sec x\sec(a-x)$ is symmetrical about $x = \frac{1}{2}a$ and therefore, by 9.3, $\int_0^a x\sec x\sec(a-x)\,dx = a\operatorname{cosec} a\log\sec a$.

9.32. Write $5-x = k\cos^2\theta$, $x-1 = k\sin^2\theta$, $0 < \theta < \frac{1}{2}\pi$, then, adding, we find $k = 4$; when $x = 2$, $\sin\theta = \frac{1}{2}$ and so $\theta = \frac{1}{6}\pi$, and when $x = 3$, $\sin\theta = 1/\sqrt{2}$, $\theta = \frac{1}{4}\pi$. Integral becomes

$$2\int_{\frac{1}{6}\pi}^{\frac{1}{4}\pi} 2\sin\theta\cos\theta\,d\theta/(\sin\theta+\cos\theta)$$

$$= 2\int_{\frac{1}{6}\pi}^{\frac{1}{4}\pi} \{(\sin\theta+\cos\theta)^2-1\}\,d\theta/(\sin\theta+\cos\theta)$$

$$= 2\int_{\frac{1}{6}\pi}^{\frac{1}{4}\pi} (\sin\theta+\cos\theta)\,d\theta - \sqrt{2}\int_{\frac{1}{6}\pi}^{\frac{1}{4}\pi} d\theta/\sin(\theta+\tfrac{1}{4}\pi)$$

$$= \sqrt{3}-1-\sqrt{2}\int_{5\pi/12}^{\frac{1}{2}\pi} d\phi/\sin\phi, \quad \phi = \theta+\tfrac{1}{4}\pi,$$

$$= \sqrt{3}-1+\sqrt{2}\log\tan 5\pi/24.$$

9.33. We must show that $\int_{1/n}^{\pi} \frac{\sin x}{x}\,dx$ is convergent; let $N > n$ then

$$0 < \int_{1/N}^{1/n} \frac{\sin x}{x}\,dx < \int_{1/N}^{1/n} dx = (1/n-1/N) < 1/n \quad (\text{since } \sin x < x, \ x > 0),$$

which proves that $\int_{1/n}^{\pi} \frac{\sin x}{x}\,dx$ converges.

9.34. Write $x = \tan t$, then integral becomes $\int_0^{\frac{1}{4}\pi} \log(1+\tan t)\,dt$, since $dx/dt = \sec^2 t = 1+x^2$. But

$$1+\tan t = (\sin t+\cos t)/\cos t = \sqrt{2}\cos(t-\tfrac{1}{4}\pi)/\cos t$$

and the integral equals

$$\int_0^{\frac{1}{4}\pi} \log\sqrt{2}\, dt + \int_0^{\frac{1}{4}\pi} \log\cos(\tfrac{1}{4}\pi - t)\, dt - \int_0^{\frac{1}{4}\pi} \log\cos t\, dt$$

$$= \tfrac{1}{8}\pi \log 2 - \int_{\frac{1}{4}\pi}^{0} \log\cos u\, du - \int_0^{\frac{1}{4}\pi} \log\cos t\, dt, \quad \text{writing } u = \tfrac{1}{4}\pi - t,$$

$$= \tfrac{1}{8}\pi \log 2.$$

9.35. Since $\int_a^\infty \phi(x)\, dx$ exists we can determine N so that $\int_N^n \phi(x)\, dx < 1/k$, $n > N$; but in (N, n), $\phi(x) \geqslant \phi(n)$, and so $(n-N)\phi(n) < 1/k$, i.e.

$$n\phi(n) < n/(n-N)k < 2/k,$$

if $n > 2N$, and so $n\phi(n) \to 0$.

9.36. $\int_0^1 f(xt)\, dt = \int_0^x f(u)\frac{du}{x}$, so that $\int_0^x f(u)\, du = 0$ for all x, and therefore $D_x \int_0^x f(u)\, du = 0$, i.e. $f(x) = 0$.

9.4.
$$\int_0^1 x^{n-1}\, dx/(1+\sqrt{x}) < \int_0^1 x^{n-1}\, dx = 1/n;$$

$$\int_0^1 x^{n-1}\, dx/(1+\sqrt{x}) > \int_0^1 x^{n-1}(1-\sqrt{x})\, dx = 1/n - 2/(2n+1);$$

$$\tfrac{1}{2}\int_0^1 x^{n-1}\, dx/(1+\sqrt{x}) = \int_0^1 y^{2n-1}\, dy/(1+y), \quad y = \sqrt{x},$$

$$= \int_0^1 \{(1+y^{2n-1})/(1+y)\}\, dy - \int_0^1 dy/(1+y)$$

$$= \int_0^1 (1-y+y^2-\dots+y^{2n-2})\, dy - \int_0^1 dy/(1+y)$$

$$= 1-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\dots+1/(2n-1)-\log 2,$$

whence

$$1/2n - 1/(2n+1) < 1-\tfrac{1}{2}+\tfrac{1}{3}-\dots+1/(2n-1)-\log 2 < 1/2n$$

and so $1-\frac{1}{2}+\frac{1}{3}-\dots = \log 2$.

9.401.
$$\int_A^B (\sin x/x)\, dx = [-\cos x/x]_A^B - \int_A^B (\cos x/x^2)\, dx;$$

but

$$\left|\int_A^B (\cos x/x^2)\, dx\right| < \int_A^B |\cos x/x^2|\, dx < \int_A^B (1/x^2)\, dx, \quad \text{since} \quad |\cos x| < 1,$$

$$= 1/A - 1/B.$$

Hence
$$\left|\int_A^B (\sin x/x)\, dx\right| \leqslant |\cos A/A - \cos B/B| + 1/A - 1/B$$
$$\leqslant 1/A + 1/B + 1/A - 1/B = 2/A.$$
In particular $\left|\int_n^N (\sin x/x)\, dx\right| < 2/n$ so that $\int_a^x (\sin x/x)\, dx$ is convergent, and $\int_a^\infty (\sin x/x)\, dx$ exists.

9.41. $$l(x^{-1}) = \int_1^{1/x} (1/t)\, dt = \int_1^x u(-u^{-2})\, du, \quad u = 1/t,$$
$$= -\int_1^x (1/u)\, du = -l(x).$$
$$l(xy) = \int_1^{xy} (1/t)\, dt = \int_1^x (1/t)\, dt + \int_x^{xy} (1/t)\, dt$$
$$= \int_1^x (1/t)\, dt + \int_1^y (1/xu)x\, du, \quad t = xu,$$
$$= l(x) + l(y).$$
$$l(x^y) = \int_1^{x^y} (1/t)\, dt = \int_1^x (1/u^y) yu^{y-1}\, du, \quad t = u^y,$$
$$= y\int_1^x (1/u)\, du = yl(x).$$

$$y = (1 + x + x^2/2! + \ldots + x^k/k!)e^{-x},$$
$$\frac{dy}{dx} = -(x^k/k!)e^{-x} = \sum (-1)^{p+1} x^{k+p}/(k!)(p!)$$
whence the result follows by integrating.

9.51. $\operatorname{sh}\alpha/e^\beta - \operatorname{sh} 3\alpha/3e^{3\beta} + \ldots$
$$= \tfrac{1}{2}(e^{\alpha-\beta} - e^{3(\alpha-\beta)}/3 + e^{5(\alpha-\beta)}/5 - \ldots) - \tfrac{1}{2}(e^{-(\alpha+\beta)} - e^{-3(\alpha+\beta)}/3 + e^{-5(\alpha+\beta)}/5 - \ldots)$$
$= \tfrac{1}{2}[\tan^{-1}e^{(\alpha-\beta)} - \tan^{-1}e^{-(\alpha+\beta)}]$ provided $e^{\alpha-\beta} \leqslant 1$, $e^{-(\alpha+\beta)} \leqslant 1$,

i.e. $-\beta \leqslant \alpha \leqslant \beta$,
$$= \tfrac{1}{2}\tan^{-1}[(e^{\alpha-\beta} - e^{-\alpha-\beta})/(1 + e^{-2\beta})] = \tfrac{1}{2}\tan^{-1}[(e^\alpha - e^{-\alpha})/(e^\beta + e^{-\beta})]$$
$$= \tfrac{1}{2}\tan^{-1}\{\operatorname{sh}\alpha/\operatorname{ch}\beta\}.$$

9.52. Since $(x-\alpha)(x-\beta) = x^2 - ax + b$, replacing x by $-1/x$, we have $(1+\alpha x)(1+\beta x) = 1 + ax + bx^2$ and so
$$\log(1 + ax + bx^2) = \log(1+\alpha x) + \log(1+\beta x)$$
$$= (\alpha+\beta)x - \tfrac{1}{2}(\alpha^2+\beta^2)x^2 + \tfrac{1}{3}(\alpha^3+\beta^3)x^3 - \ldots,$$
provided $-1 < \alpha x \leqslant 1$, $-1 < \beta x \leqslant 1$, i.e. if $|\alpha| \geqslant |\beta|$ then $x^2 < 1/\alpha^2$ or $x = 1/\alpha$, and if $|\alpha| < |\beta|$ then $x^2 < 1/\beta^2$ or $x = 1/\beta$.

Take $\alpha = e^{\theta}$, $\beta = e^{-\theta}$ then $(\alpha^n+\beta^n) = 2\operatorname{ch} n\theta$ and so

$$x\operatorname{ch}\theta-\tfrac{1}{2}x^2\operatorname{ch}2\theta+\tfrac{1}{3}x^3\operatorname{ch}3\theta-\ldots = \tfrac{1}{2}\log(1+xe^{\theta})(1+xe^{-\theta})$$
$$= \tfrac{1}{2}\log(1+2x\operatorname{ch}\theta+x^2)$$

provided $-e^{-|\theta|} < x \leqslant e^{-|\theta|}$.

9.53. Since $\cos(n+2)\theta+\cos n\theta = 2\cos\theta\cos(n+1)\theta$, therefore by Example 1.8

$$\sum_{n\geq 1} x^{n-1}\cos n\theta = \{\cos\theta+(\cos 2\theta-2\cos^2\theta)x\}/(1-2x\cos\theta+x^2)$$

and so, changing the sign of x,

$$\cos\theta-x\cos 2\theta+x^2\cos 3\theta-\ldots = (x+\cos\theta)/(1+2x\cos\theta+x^2)$$

the series being absolutely convergent for $|x| < 1$. Hence by integrating from 0 to x, $|x| < 1$,

$$x\cos\theta-\tfrac{1}{2}x^2\cos 2\theta+\tfrac{1}{3}x^3\cos 3\theta-\ldots = \tfrac{1}{2}\log(1+2x\cos\theta+x^2).$$

Similarly, since $\sin(2n+5)\theta+\sin(2n+1)\theta = 2\sin(2n+3)\theta\cos 2\theta$, therefore

$$\sin\theta+t\sin 3\theta+t^2\sin 5\theta+t^3\sin 7\theta+\ldots = \frac{\sin\theta+(\sin 3\theta-2\sin\theta\cos 2\theta)t}{1-2t\cos 2\theta+t^2}$$
$$= (1+t)\sin\theta/(1-2t\cos 2\theta+t^2), \quad |t| < 1.$$

Hence replacing t by x^2 and integrating we have

$$x\sin\theta+\tfrac{1}{3}x^3\sin 3\theta+\tfrac{1}{5}x^5\sin 5\theta+\ldots = \int_0^x (1+x^2)\sin\theta\,dx/(1-2x^2\cos 2\theta+x^4)$$
$$= \frac{1}{2}\int_0^x \left(\frac{\sin\theta}{1-2x\cos\theta+x^2}+\frac{\sin\theta}{1+2x\cos\theta+x^2}\right)dx$$
$$= \tfrac{1}{2}[\tan^{-1}\{(x-\cos\theta)/\sin\theta\}+\tan^{-1}\{(x+\cos\theta)/\sin\theta\}], \quad \text{by 9.2,}$$
$$= \tfrac{1}{2}\tan^{-1}\{2x\sin\theta/(1-x^2)\}, \quad |x| < 1.$$

9.531. By Example 1.8,

$$\cos\theta+r\cos 2\theta+r^2\cos 3\theta+\ldots = \frac{\cos\theta-r}{1-2r\cos\theta+r^2}, \quad |r| < 1,$$

$$\sin\theta+r\sin 2\theta+r^2\sin 3\theta+\ldots = \frac{\sin\theta}{1-2r\cos\theta+r^2}, \quad |r| < 1,$$

and so

$$\sum \frac{r^n}{n}\cos n\theta = \int_0^r \frac{\cos\theta-r}{1-2r\cos\theta+r^2}\,dr = -\tfrac{1}{2}\log(1-2r\cos\theta+r^2),$$

$$\sum \frac{r^n}{n}\sin n\theta = \int_0^r \frac{\sin\theta}{1-2r\cos\theta+r^2}\,dr = \int_0^r \frac{\sin\theta}{(r-\cos\theta)^2+\sin^2\theta}\,dr$$
$$= \tan^{-1}\left(\frac{r-\cos\theta}{\sin\theta}\right)+\tan^{-1}\left(\frac{\cos\theta}{\sin\theta}\right)$$
$$= \tan^{-1}\frac{r\sin\theta}{1-r\cos\theta}.$$

9.532. Write $C = C(r) = \sum_{n\geqslant 0} \frac{r^n}{n!}\cos n\theta$, $S = S(r) = \sum_{n\geqslant 0} \frac{r^n}{n!}\sin n\theta$; since $|\cos n\theta| \leqslant 1$, $|\sin n\theta| \leqslant 1$, therefore the two series are absolutely convergent for all values of r and θ.

$$C'(r) = \sum \frac{r^n}{n!}\cos(n+1)\theta = C\cos\theta - S\sin\theta,$$

$$S'(r) = \sum \frac{r^n}{n!}\sin(n+1)\theta = C\sin\theta + S\cos\theta.$$

Write $c = c(r) = \cos(r\sin\theta)$, $s = s(r) = \sin(r\sin\theta)$ so that $D_r c = -s\sin\theta$, $D_r s = c\sin\theta$. Then if $J = J(r) = Cc+Ss$ and $H = H(r) = Cs-Sc$ we have

$$\begin{aligned} J'(r) &= (C\cos\theta - S\sin\theta)c - Cs\sin\theta + (C\sin\theta + S\cos\theta)s + Sc\sin\theta \\ &= \cos\theta(Cc+Ss) = J\cos\theta. \end{aligned}$$

Hence $D_r(Je^{-r\cos\theta}) = 0$ and so

$$Je^{-r\cos\theta} = \text{constant} = J(0) = 1, \quad \text{i.e.} \quad J = e^{r\cos\theta}.$$

Similarly $\quad He^{-r\cos\theta} = \text{constant} = H(0) = 0, \quad \text{i.e.} \quad H = 0.$

Therefore $\quad e^{r\cos\theta}\cos(r\sin\theta) = Jc+Hs = C(c^2+s^2) = C$

and $\quad e^{r\cos\theta}\sin(r\sin\theta) = Js-Hc = S(s^2+c^2) = S.$

9.54. Let $F(x) = \int_a^x \phi(x)\,dx$ then $F(n+1)-F(n) = \int_n^{n+1} \phi(x)\,dx$; but since $\phi(x)$ steadily decreases, $\phi(n+1) < \int_n^{n+1} \phi(x)\,dx < \phi(n)$ and so

$$\phi(n+1) < F(n+1)-F(n) < \phi(n).$$

Thus $\sum \phi(n)$ converges and diverges together with $\sum \{F(n+1)-F(n)\}$; but

$$\sum_a^{n-1}\{F(r+1)-F(r)\} = F(n) = \int_a^n \phi(x)\,dx;$$

$$\int_0^n x^k e^{-x}\,dx = -(k!)[e^{-x}]_0^n = (k!)(1-e^{-n}) \to k!,$$

and so $\sum n^k e^{-n}$ converges.

9.6. We have

$$\tfrac{1}{2}+\cos x+\cos 2x+\ldots+\cos nx = \sin(n+\tfrac{1}{2})x/2\sin\tfrac{1}{2}x, \quad 0 < x < 2\pi;$$

let x lie in $[0, 2\pi]$ and integrate from π to x, then

$$\begin{aligned} &\left|\tfrac{1}{2}(x-\pi)+\sin x+\tfrac{1}{2}\sin 2x+\tfrac{1}{3}\sin 3x+\ldots+\frac{1}{n}\sin nx\right| \\ &\qquad = \left|\int_\pi^x \sin(n+\tfrac{1}{2})x\,dx/2\sin\tfrac{1}{2}x\right| \\ &\qquad = \left|\left[\frac{1}{2n+1}\,\frac{\cos(n+\frac{1}{2})x}{\sin\frac{1}{2}x}\right]_\pi^x + \frac{1}{2(2n+1)}\int_\pi^x \frac{\cos(n+\frac{1}{2})x\cos\frac{1}{2}x}{\sin^2\frac{1}{2}x}\,dx\right|, \end{aligned}$$

integrating by part

$$< \frac{1}{2n+1}\{\operatorname{cosec}\tfrac{1}{2}x+\tfrac{1}{2}\pi\operatorname{cosec}^2\tfrac{1}{2}x\} \to 0 \quad \text{for a fixed } x.$$

The sum of the series $\sin x+\frac{\sin 2x}{2}+\frac{\sin 3x}{3}+\ldots$ may also be derived from Example 9.531. For $\sum \frac{r^n}{n}\sin nx = \tan^{-1}\left(\frac{r\sin x}{1-r\cos x}\right)$, $|r|<1$; the right-hand side of this equation is continuous for $0 \leqslant r \leqslant 1$, when $0<x<2\pi$, and since $\left|\sum_1^n \sin rx\right| < 1/|\sin\frac{1}{2}x|$, and $1/n \to 0$, therefore by Theorem 1.913 the series $\sum \frac{\sin nx}{n}$ is convergent. Hence

$$\sum \frac{\sin nx}{n} = \tan^{-1}\frac{\sin x}{1-\cos x} = \tan^{-1}(\cot\tfrac{1}{2}x) = \tan^{-1}\{\tan(\tfrac{1}{2}\pi-\tfrac{1}{2}x)\} = \tfrac{1}{2}\pi-\tfrac{1}{2}x$$

since $-\frac{1}{2}\pi < \frac{1}{2}\pi-\frac{1}{2}x < \frac{1}{2}\pi$ when $0<x<2\pi$.

Similarly from Example 9.531 we obtain

$$\sum \frac{\cos nx}{n} = -\tfrac{1}{2}\log(2-2\cos x) = -\log(2\sin\tfrac{1}{2}x), \quad 0<x<2\pi.$$

9.61. If $\phi(t) = \frac{1}{2}(t-\pi)+\sin t+\frac{\sin 2t}{2}+\ldots+\frac{\sin nt}{n}$ then by the first proof of 9.6,

$$|\phi(t)| \leqslant (\operatorname{cosec}\tfrac{1}{2}t+\tfrac{1}{2}\pi\operatorname{cosec}^2\tfrac{1}{2}t)/(2n+1) = k_t/(2n+1), \quad \text{say}$$

hence, integrating from x to π, $0<x<2\pi$

$$\left|\left[\frac{(t-\pi)^2}{4}-\cos t-\frac{\cos 2t}{2^2}-\frac{\cos 3t}{3^2}-\ldots-\frac{\cos nt}{n^2}\right]_x^\pi\right| = \left|\int_x^\pi \phi(t)\,dt\right| < \pi k_x/(2n+1)$$

which tends to zero for a fixed x. Thus

$$\sum \frac{\cos nx}{n^2} = \frac{(x-\pi)^2}{4}-\sum\frac{(-1)^{n+1}}{n^2} = \frac{(x-\pi)^2}{4}-\sigma, \quad \text{say}.$$

This equation holds for $0<x<2\pi$; since both sides of the equation are continuous in any interval (Example 5.7), the equality holds also for $x=0$ and $x=2\pi$. Taking $x=0$ we find

$$\sum\frac{1}{n^2}+\sum\frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{4}, \quad \text{i.e.} \quad \sum\frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

Since $\quad \sum\frac{1}{n^2} = \sum\frac{1}{(2n+1)^2}+\frac{1}{4}\sum\frac{1}{n^2}$, therefore $\sum\frac{1}{n^2} = \frac{\pi^2}{6}$,

and $$\sigma = \sum\frac{(-1)^{n+1}}{n^2} = \sum\frac{1}{(2n+1)^2}-\frac{1}{4}\sum\frac{1}{n^2} = \frac{\pi^2}{12}.$$

Thus $$\sum\frac{\cos nx}{n^2} = \frac{x^2}{4}-\frac{\pi x}{2}+\frac{\pi^2}{6}, \quad 0 \leqslant x \leqslant 2\pi.$$

9.62. $$\int_0^c \frac{\phi(x)}{\sqrt{(cx-x^2)}}\,dx = 2\int_0^{\frac{1}{2}\pi}\phi(c\sin^2\theta)\,d\theta, \quad x = c\sin^2\theta;$$

but $\phi(x) = \sum_0^n a_n x^r$ and so

$$\int_0^c \frac{\phi(x)}{\sqrt{(cx-x^2)}}\,dx = 2\sum_0^n a_r c^r \int_0^{\frac{1}{2}\pi} \sin^{2r}\theta\,d\theta$$

which is a polynomial in c of the nth degree.

9.63. If $F(x) = \int_a^x \phi(y)\,dy \int_x^b \psi(y)\,dy$ then $F(a) = F(b) = 0$ and

$$F'(x) = \phi(x)\int_x^b \psi(y)\,dy - \psi(x)\int_a^x \phi(y)\,dy$$

whence integrating from a to b, the result follows.

9.64. $$\int_0^{x+a} \phi(t)\,dt = \int_0^a + \int_a^{a+x} \phi(t)\,dt = \int_0^a \phi(t)\,dt + \int_0^x \phi(u)\,du, \quad t = a+u,$$

since $\phi(x)$ has period a,

$$= \int_0^x \phi(u)\,du,$$

for

$$\int_0^a \phi(t)\,dt = \int_0^a f(t)\,dt - \frac{1}{a}\int_0^a \left\{\int_0^a f(x)\,dx\right\}dt = \int_0^a f(t)\,dt - \int_0^a f(t)\,dt = 0.$$

9.7. In $[0, \frac{1}{2}\pi]$, $0 < \sin x < 1$, and so $\sin^{2n+1}x < \sin^{2n}x < \sin^{2n-1}x$, whence

$$\int_0^{\frac{1}{2}\pi} \sin^{2n+1}x\,dx < \int_0^{\frac{1}{2}\pi} \sin^{2n}x\,dx < \int_0^{\frac{1}{2}\pi} \sin^{2n-1}x\,dx,$$

i.e. $\qquad 2n!/(2n+1)! < \{(2n-1)!/2n!\}\frac{1}{2}\pi < (2n-2)!/(2n-1)!.$

Denote $\{2n!/(2n-1)!\}^2(2/\pi)-2n$ by θ_n, so that $0 < \theta_n < 1$; therefore

$$\tfrac{1}{2}\pi = \{2n!/(2n-1)!\}^2/(2n+\theta_n)$$

for all n, i.e.

$$\{(2n+\theta_n)/(2n+1)\}\tfrac{1}{2}\pi = \{2n!/(2n-1)!\}^2/(2n+1).$$

Since $0 < \theta_n < 1$, therefore $(2n+\theta_n)/(2n+1) \to 1$ and so

$$\{2n!/(2n-1)!\}^2/(2n+1) \to \tfrac{1}{2}\pi,$$

i.e. $$\tfrac{1}{2}\pi = \frac{2.2}{1.3}\cdot\frac{4.4}{3.5}\cdot\frac{6.6}{5.7}\cdots\frac{2n.2n}{2n-1.2n+1}\cdots.$$

Since $2n/(2n+1) \to 1$, it follows also that $2n!/(2n-1)!\sqrt{n} \to \sqrt{\pi}$.

9.8. $\int_m^n e^{-x^2}\,dx < \int_m^n (1/x^2)\,dx = 1/m - 1/n < 1/m$ and so $\int_0^\infty e^{-x^2}\,dx$ converges. From Example 4.9,

$$1-x^2 < e^{-x^2} < 1/(1+x^2) \quad \text{and so} \quad (1-x^2)^k < e^{-kx^2} < 1/(1+x^2)^k,$$

whence

$$\int_0^\infty e^{-x^2}\,dx > \int_0^n e^{-x^2}\,dx = \sqrt{k}\int_0^{n/\sqrt{k}} e^{-kx^2}\,dx > \sqrt{k}\int_0^1 e^{-kx^2}\,dx > \sqrt{k}\int_0^1 (1-x^2)^k\,dx$$
$$= \{2k!/(2k+1)!\}\sqrt{k}$$

Furthermore,

$$\int_0^n e^{-x^2}\,dx = \sqrt{k}\int_0^{n/\sqrt{k}} e^{-kx^2}\,dx < \sqrt{k}\int_0^{n/\sqrt{k}} \{1/(1+x^2)^k\}\,dx < \sqrt{k}\int_0^\infty \{1/(1+x^2)^k\}\,dx$$
$$= \{(2k-3)!/(2k-2)!\}\left(\frac{\pi\sqrt{k}}{2}\right)$$

and so $$\int_0^\infty e^{-x^2}\,dx \leqslant \{(2k-3)!/(2k-2)!\}(\pi\sqrt{k}/2);$$

thus $$2k!\sqrt{k}/(2k+1)! < \int_0^\infty e^{-x^2}\,dx \leqslant (\tfrac{1}{2}\sqrt{k}\pi)(2k-3)!/(2k-2)!,$$

whence

$$\frac{k}{2k+1}\left[\frac{2k!}{(2k-1)!}\cdot\frac{1}{\sqrt{k}}\right] < \int_0^\infty e^{-x^2}\,dx < \tfrac{1}{2}\pi\frac{2k}{2k-1}\left[\frac{(2k-1)!}{(2k)!}\sqrt{k}\right].$$

But $$\frac{k}{2k+1}\,\frac{2k!}{(2k-1)!}\cdot\frac{1}{\sqrt{k}} \to \tfrac{1}{2}\sqrt{\pi} \quad \text{by 9.7},$$

and $$\tfrac{1}{2}\pi\frac{2k}{2k-1}\,\frac{(2k-1)!}{2k!}\sqrt{k} \to \tfrac{1}{2}\pi\frac{1}{\sqrt{\pi}} = \tfrac{1}{2}\sqrt{\pi}$$

so that $\int_0^\infty e^{-x^2}\,dx = \tfrac{1}{2}\sqrt{\pi}$.

9.9. In the interval $(-1,0)$, $x = -\sqrt{y}$ and in $(0,2)$, $x = +\sqrt{y}$, hence

$$\int_{-1}^2 x\,dx/x^{\frac{4}{3}} = \int_{-1}^0 x\,dx/x^{\frac{4}{3}} + \int_0^2 x\,dx/x^{\frac{4}{3}} = \lim\int_{-1}^{-1/n} x\,dx/x^{\frac{4}{3}} + \lim\int_{1/n}^2 x\,dx/x^{\frac{4}{3}}$$
$$= \lim\int_{1/n^2}^1 dy/2y^{\frac{2}{3}} + \lim\int_{1/n^2}^4 dy/2y^{\frac{2}{3}} = 3[y^{\frac{1}{3}}]_{1/n^2}^1 + 3[y^{\frac{1}{3}}]_{1/n^2}^4 \to 3(1+\sqrt[3]{2}).$$

If $y = x^2-6x+13$ then $(x-3)^2 = y-4$ and so $x-3 = -\sqrt{(y-4)}$ when x lies in $(1,3)$ and $x-3 = +\sqrt{(y-4)}$ when x lies in $(3,7)$. Hence $dx/dy = -1/2\sqrt{(y-4)}$ when x lies in $(1,3)$ and y lies in $(4,8)$, and $dx/dy = +1/2\sqrt{(y-4)}$ when x lies in $(3,7)$ and y lies in $(4,20)$. Hence

$$\int_1^7 (x^2-6x+13)\,dx = \int_1^3 (x^2-6x+13)\,dx + \int_3^7 (x^2-6x+13)\,dx$$
$$= \int_8^4 \frac{y}{2\sqrt{(y-4)}}\,dy + \int_4^{20} \frac{y}{2\sqrt{(y-4)}}\,dy$$
$$= \int_4^8 \sqrt{(y-4)}\,dy + 4\int_4^8 dy/\sqrt{(y-4)} + \tfrac{1}{2}\int_8^{20}\sqrt{(y-4)}\,dy + 2\int_8^{20} dy/\sqrt{(y-4)}$$
$$= \tfrac{2}{3}[(y-4)^{\frac{3}{2}}]_4^8 + 8[(y-4)^{\frac{1}{2}}]_4^8 + \tfrac{1}{3}[(y-4)^{\frac{3}{2}}]_8^{20} + 4[(y-4)^{\frac{1}{2}}]_8^{20}$$
$$= \tfrac{16}{3} + 16 + \tfrac{64}{3} - \tfrac{8}{3} + 16 - 8 = 48.$$

9.91. We prove first that the sequence $1+\frac{1}{2}+\frac{1}{3}+...+1/(n-1)-\log n$ is convergent; for $0 < \int_r^{r+1} (1/r-1/x)\,dx < 1/r-1/(r+1)$, and $\sum \{1/r-1/(r+1)\}$ converges so that $\sum_{r=1}^{n-1} \int_r^{r+1} (1/r-1/x)\,dx = \sum_1^{n-1} 1/r - \int_1^n (1/x)\,dx = \sum_1^{n-1} 1/r - \log n$ converges, to γ (say). Next we prove that $\sum [r\log(1+1/r)-1+1/2r]$ converges; if $0 < x < 1$ the expansion of $\log(1+x)$ is an alternating series of steadily decreasing terms, so that $x-x^2/2 < \log(1+x) < x-x^2/2+x^3/3$ and therefore

$$0 < r\log(1+1/r)-1+1/2r < 1/3r^2.$$

Since $\sum 1/r^2$ converges, therefore the positive series

$$\sum [r\log(1+1/r)-1+1/2r]$$

also converges, to δ (say).

We have

$$\begin{aligned}\log n! &= \sum_1^n \log r \\ &= \sum_1^{n-1}\Big[\log(r+1) - \int_r^{r+1} \log x\,dx\Big] + \int_1^n \log x\,dx \\ &= \sum_1^{n-1} [1/2r - \{r\log(1+1/r)-1+1/2r\}] + n\log n-(n-1) \\ &= \tfrac{1}{2}\Big(\sum_1^{n-1} 1/r - \log n\Big) - \sum_1^{n-1}\{r\log(1+1/r)-1+1/2r\} + \\ &\qquad\qquad + 1 + \tfrac{1}{2}\log n + n\log n - n.\end{aligned}$$

And so $\qquad \log n! - \log\{(n/e)^n\sqrt{n}\} \to \gamma/2-\delta+1,$

whence $\qquad n!/(n/e)^n\sqrt{n} \to e^{\gamma/2-\delta+1} = C,\quad$ say.

The value of C may readily be determined from Wallis's formula (Ex. 9.7). For

$$\begin{aligned}\tfrac{1}{2}\pi = \lim\Big\{\frac{2n!}{(2n-1)!}\Big\}^2\frac{1}{2n+1} &= \lim\frac{(2n!)^4}{\{2n!(2n-1)!\}^2(2n+1)} = \lim\frac{2^{4n}(n!)^4}{(2n!)^2(2n+1)} \\ &= \lim\frac{[n!/(n/e)^n\sqrt{n}]^4 n/2}{[2n!/(2n/e)^{2n}\sqrt{2n}]^2(2n+1)} = \frac{C^2}{4}.\end{aligned}$$

Hence $C = \sqrt{(2\pi)}$, since C is positive, and so

$$n!/(n/e)^n\sqrt{(2\pi n)} \to 1.$$

Thus for large values of n an approximate value of $n!$ is $(n/e)^n\sqrt{(2\pi n)}$.

9.92.
$$\begin{aligned}\int_{1/n}^{1} \frac{dx}{\sqrt{(\log 1/x)}} &= -\int_{\sqrt{(\log n)}}^{0} \frac{2y}{y} e^{-y^2}\,dy, \quad x = e^{-y^2}, \\ &= 2\int_0^{\sqrt{(\log n)}} e^{-y^2}\,dy \to 2\int_0^\infty e^{-y^2}\,dy = \sqrt{\pi},\end{aligned}$$

by Examples 9.8 and 1.9, since $\int_0^a e^{-y^2}\,dy$ is a strictly increasing function.

9.93. In $(0,1]$, $1+t \leqslant 1/(1-t)$, whence $\log(1-t) \leqslant -(t+\frac{1}{2}t^2)$ and so $e^{nt}(1-t)^n \leqslant e^{-\frac{1}{2}nt^2}$. Hence

$$\sqrt{n}\int_0^1 e^{nt}(1-t)^n\,dt \leqslant \sqrt{n}\int_0^1 e^{-\frac{1}{2}nt^2}\,dt = \sqrt{2}\int_0^{\sqrt{(\frac{1}{2}n)}} e^{-u^2}\,du, \quad t = u\sqrt{(2/n)},$$

$$\leqslant \sqrt{2}\int_0^\infty e^{-u^2}\,du = \sqrt{(\tfrac{1}{2}\pi)}.$$

In $(0, 1-1/k)$,

$$-\log(1-t) = \int_0^t \frac{1}{1-t}\,dt = \int_0^t \left(1+t+\frac{t^2}{1-t}\right)dt \leqslant t+\tfrac{1}{2}t^2+\tfrac{1}{3}kt^3$$

since $\dfrac{1}{1-t} < k$ and therefore

$$e^{nt}(1-t)^n \geqslant e^{-\frac{1}{2}nt^2-\frac{1}{3}knt^3} \geqslant e^{-\frac{1}{2}nt^2}(1-\tfrac{1}{3}knt^3)$$

from which

$$\sqrt{(\tfrac{1}{2}\pi)} \geqslant \sqrt{n}\int_0^1 e^{nt}(1-t)^n\,dt \geqslant \sqrt{n}\int_0^{1-1/k} e^{nt}(1-t)^n\,dt$$

$$\geqslant \sqrt{n}\int_0^{1-1/k} e^{-\frac{1}{2}nt^2}(1-\tfrac{1}{3}knt^3)\,dt$$

$$= \sqrt{2}\int_0^{(1-1/k)\sqrt{(\frac{1}{2}n)}} e^{-u^2}\,du - \frac{2}{3}\frac{k}{\sqrt{n}}\int^{\frac{1}{2}n(1-1/k)^2} e^{-v}v\,dv,$$

$$u = t\sqrt{(\tfrac{1}{2}n)}, \quad v = \tfrac{1}{2}nt^2,$$

$$\to \sqrt{(\tfrac{1}{2}\pi)},$$

taking $k = n^{\frac{1}{4}}$, since $\int^N e^{-v}v\,dv \to 1$ and $n^{\frac{1}{4}}/n^{\frac{1}{2}} \to 0$, i.e.

$$\sqrt{n}\int_0^1 e^{nt}(1-t)^n\,dt \to \sqrt{(\tfrac{1}{2}\pi)}.$$

Accordingly

$$\int_0^n e^{-x}\frac{x^n}{n!}\,dx = \frac{n^{n+1}}{n!}\int_0^1 e^{-ny}y^n\,dy, \quad x = ny,$$

$$= e^{-n}\frac{n^{n+1}}{n!}\int_0^1 e^{nt}(1-t)^n\,dt, \quad y = 1-t,$$

$$= \frac{\sqrt{(2\pi n)}\,.\,n^n}{e^n n!}\cdot\frac{\sqrt{n}}{\sqrt{(2\pi)}}\int_0^1 e^{nt}(1-t)^n\,dt$$

$$\to \tfrac{1}{2}, \quad \text{using Example 9.91.}$$

But
$$\int_0^n e^{-x}\frac{x^n}{n!}\,dx = 1-e^{-n}\left\{1+n+\frac{n^2}{2}+\dots+\frac{n^n}{n!}\right\},$$

whence
$$e^{-n}\left\{1+n+\dots+\frac{n^n}{n!}\right\} \to \tfrac{1}{2}.$$

9.94. If $a > 0$, $\int_{1/N}^{1/n} t^{a-1}e^{-t}\,dt < \int_{1/N}^{1/n} t^{a-1}\,dt = [t^a/a]_{1/N}^{1/n} < 1/an^a \to 0$, and so $\int_0^k t^{a-1}e^{-t}\,dt$ exists for any k. Furthermore, if ν is an integer greater than $a+1$ then $t^a e^{-t} < \nu!/t^{\nu-a}$ so that $[t^a e^{-t}]_0^n \to 0$ and

$$\int_n^N t^a e^{-t}\,dt < \nu!/\{\nu-(a+1)\}n^{\nu-a-1} \to 0.$$

Thus $\int_0^\infty t^a e^{-t}\,dt$ exists; but

$$\int_0^n t^{a-1}e^{-t}\,dt = \left[\frac{t^a}{a}e^{-t}\right]_0^n + \frac{1}{a}\int_0^n t^a e^{-t}\,dt$$

and therefore $\int_0^\infty t^{a-1}e^{-t}\,dt = \dfrac{1}{a}\int_0^\infty t^a e^{-t}\,dt$, i.e. $\Gamma(a)$ is defined for $a > 0$, and $\Gamma(a+1) = a\Gamma(a)$.

$$\Gamma(\tfrac{1}{2}) = \lim \int_0^n t^{-\frac{1}{2}}e^{-t}\,dt = \lim \int_0^{\sqrt{n}} u^{-1}.e^{-u^2}.2u\,du = 2\int_0^\infty e^{-u^2}\,du = \sqrt{\pi}.$$

If m is even

$$\Gamma\left(\frac{m+1}{2}\right) = \frac{m-1}{2}\Gamma\left(\frac{m-1}{2}\right) = \frac{m-1}{2}.\frac{m-3}{2}\Gamma\left(\frac{m-3}{2}\right)$$

$$= \frac{m-1}{2}.\frac{m-3}{2}\cdots\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = (m-1)!\sqrt{\pi}/2^{\frac{1}{2}m},$$

and if m is odd

$$\Gamma\left(\frac{m+1}{2}\right) = \frac{m-1}{2}.\frac{m-3}{2}\cdots\frac{2}{2}\Gamma(1) = (m-1)!/2^{\frac{1}{2}(m-1)}.$$

Thus, if m and n are both even,

$$\frac{1}{2}\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)\Big/\Gamma\left(\frac{m+n+2}{2}\right) = \frac{\frac{1}{2}.(m-1)!.(n-1)!\pi/2^{\frac{1}{2}(m+n)}}{(m+n)!/2^{\frac{1}{2}(m+n)}}$$

$$= \frac{(m-1)!.(n-1)!}{(m+n)!}\tfrac{1}{2}\pi.$$

Similarly, if one of m and n is odd, or both are odd,

$$\frac{1}{2}\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)\Big/\Gamma\left(\frac{m+n+2}{2}\right) = \frac{(m-1)!(n-1)!}{(m+n)!},$$

whence the result follows, by the formula of § 9.31.

X

10. $dy/dx = y'/x' = \sin\theta/(1-\cos\theta)$;

subtangent $= y\,dx/dy = 2a\sin^2\tfrac{1}{2}\theta\tan\tfrac{1}{2}\theta$,

subnormal $= y\,dy/dx = a\sin\theta$ and $x+a\sin\theta = a\theta$.

10.11. Area $= \frac{1}{2}\int^{\pi} (xy'-x'y)\,dt = 2\int^{\pi} (2\sin t\cos 2t-\cos t\sin 2t)\,dt$, etc.

10.12. $$\text{Area} = \tfrac{1}{2}\int_{-1}^{1} x^2\,dt \quad (\text{since } y = tx),$$
$$= \tfrac{1}{2}\int_{-1}^{1} \{4/(1+t^2)^2-4/(1+t^2)+1\}\,dt, \quad \text{etc.}$$

10.13. $y = tx$, area $= \frac{1}{2}\int_0^{\infty} a^2t^2/(1+t^3)^2\,dt = [-a^2/6(1+t^3)]_0^{\infty} = a^2/6.$

10.14. $$\text{Area} = \tfrac{1}{2}\int^{\frac{1}{2}\pi} r^2\,d\theta$$
$$= \lim(25a^2/2)\int_0^{\frac{1}{2}\pi-1/n} \{\tan^4\theta\sec^2\theta/(1+\tan^5\theta)^2\}\,d\theta$$
$$= \lim[-25a^2/10(1+\tan^5\theta)]_0^{\frac{1}{2}\pi-1/n} = 5a^2/2.$$

10.15. $y = tx, \quad x = 3t/(t^3+2)$, etc.

10.16. $\text{Area} = \frac{1}{2}\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi} (2\cos\theta-1)^2\,d\theta$, etc.

10.2. Since $y = mx+c$ is the tangent to $y = f(x)$ at $x = x_0$ we have $f(x_0) = mx_0+c$, $f'(x_0) = m$ and so if $\phi(x) = f(x)-mx-c$ then
$$\phi(x_0) = \phi'(x_0) = 0$$
and result follows by Example 3.1.

10.21. $$f'(x) = a-12b^2/x^3, \qquad f''(x) = 36b^2/x^4,$$
therefore
$$y'' = e^{-f}f'^2-e^{-f}f'' = \frac{e^{-f}}{x^6}\{(ax^3-6bx-12b^2)(ax^3+6bx-12b^2)\}.$$

The cubics have no common factor, since $b \neq 0$, and each cubic has either one real root or three real roots of which one must be *simple*, that is each cubic has at least one non-repeated real root, and y'' changes sign, and vanishes, as x passes through a non-repeated real root of each cubic.

10.31. Area $= \frac{1}{2}\int_0^{2\pi} 3(\sin^2t\cos^4t+\cos^2t\sin^4t)\,dt = \frac{3}{8}\int_0^{2\pi} \sin^2 2t\,dt$, etc.

Length $= \int_0^{2\pi} 3(\cos^4t\sin^2t+\cos^2t\sin^4t)^{\frac{1}{2}}\,dt = 3\int_0^{\frac{1}{2}\pi}\sin 2t\,dt-3\int_{\frac{1}{2}\pi}^{\pi}\sin 2t\,dt = 6,$
since the positive root of $\sin^2t\cos^2t$ is $-\sin t\cos t$ in $(\frac{1}{2}\pi,\pi)$ and
$$\int_0^{2\pi} (\sin^2 2t)^{\frac{1}{2}}\,dt = 2\int_0^{\pi} (\sin^2 2t)^{\frac{1}{2}}\,dt.$$

10.311. Area $= \frac{1}{2}\int_0^\infty t^2 e^{-4t}\,dt$, length $= \int_0^\infty (2t^2-2t+1)e^{-2t}\,dt$.

10.32. $$r = +\sqrt{(x^2+y^2)} = +\sqrt{(1+\cos t)^2} = 1+\cos t$$

and $\cos\theta = x/r = \cos t$, whence $r = 1+\cos\theta$. Area $= \frac{1}{2}\int_0^{2\pi} (1+\cos\theta)^2\,d\theta$ and length

$$= \int^{2\pi} \{r^2+(dr/d\theta)^2\}^{\frac{1}{2}}\,d\theta = \sqrt{2}\int^{2\pi} \sqrt{(1+\cos\theta)}\,d\theta$$

$$= 2\int_0^{2\pi} (\cos^2\tfrac{1}{2}\theta)^{\frac{1}{2}}\,d\theta$$

$$= 2\int_0^{\pi} \cos\tfrac{1}{2}\theta\,d\theta - 2\int_{\pi}^{2\pi} \cos\tfrac{1}{2}\theta\,d\theta, \quad \text{etc.}$$

10.33. If $\pm\tau$ are the values of t for which $\operatorname{ch} t = 4$, then length of arc

$$= \int_{-\tau}^{\tau} (\operatorname{sh}^2 2t)^{\frac{1}{2}}\,dt = 2\int_0^{\tau} \operatorname{sh} 2t\,dt = [2\operatorname{ch}^2 t]_0^{\tau} = 2(16-1) = 30.$$

Length of arc of $y = 2\tan^2 x + \frac{1}{2}\log\cos^2 x$ is

$$\int_0^{\frac{1}{4}\pi} \{1+\tan^2 x(\sec^2 x-1)^2\}^{\frac{1}{2}}\,dx = \int_0^{\frac{1}{4}\pi} \sec x(1+4\tan^2 x)\,dx$$

$$= 4\int^{\frac{1}{4}\pi} \sec^3 x\,dx - 3\int^{\frac{1}{4}\pi} \sec x\,dx, \quad \text{etc.}$$

10.34. $$x = \tfrac{3}{2}a\cos t - a\cos^3 t, \qquad \text{area} = \frac{3a^2}{4}\int_0^{2\pi} \sin^2 t\,dt = 3\pi a^2/4$$

$$\text{length} = a\int^{2\pi} 3\{\sin^4 t\cos^2 t + \sin^2 t(\cos^2 t - \tfrac{1}{2})^2\}^{\frac{1}{2}}\,dt$$

$$= \tfrac{3}{2}a\int_0^{2\pi} \sqrt{(\sin^2 t)}\,dt = 3a\int_0^{\pi} \sin t\,dt = 6a.$$

10.35. The curves intersect where $(x^2-1)^2 = 9x^2/4$, i.e. where either $x^2-\frac{3}{2}x-1 = 0$ or $x^2+\frac{3}{2}x-1 = 0$, whence $x = \pm 2$ or $x = \pm\frac{1}{2}$. Further $(x^2-1)^2-\frac{9}{4}x^2$ is negative between -2 and $-\frac{1}{2}$, positive between $-\frac{1}{2}$ and $+\frac{1}{2}$, and negative between $+\frac{1}{2}$ and $+2$, and so the total area contained between the curves is

$$\int_{-2}^{-\frac{1}{2}} \{\tfrac{9}{4}x^2-(x^2-1)^2\}\,dx + \int_{-\frac{1}{2}}^{+\frac{1}{2}} \{(x^2-1)^2-\tfrac{9}{4}x^2\}\,dx + \int_{\frac{1}{2}}^{2} \{\tfrac{9}{4}x^2-(x^2-1)^2\}\,dx$$

$\left(\text{and is } \textit{not} \text{ equal to } \int_{-2}^{2} \{\tfrac{9}{4}x^2-(x^2-1)^2\}\,dx\right)$

10.36. Take $x = a\sin^3 t$, $y = a\cos^3 t$, then length of a quadrant is

$$\int_0^{\frac{1}{2}\pi} 3\sin t\cos t\sqrt{(a^2\sin^2 t+b^2\cos^2 t)}\,dt$$
$$= [(a^2\sin^2 t+b^2\cos^2 t)^{\frac{3}{2}}/(a^2-b^2)]_0^{\frac{1}{2}\pi} = (a^3-b^3)/(a^2-b^2).$$

10.37. $$s = \int_0^t a\{\sin^2 t+(1+\cos t)^2\}^{\frac{1}{2}}\,dt = 4a\sin\tfrac{1}{2}t$$

and $$\tan\psi = y'/x' = \sin t/(1+\cos t) = \tan\tfrac{1}{2}t,$$

whence $s = 4a\sin\psi$.

10.4. We have $r = a\,\mathrm{th}\,\tfrac{1}{2}\theta$, whence

$$s = \int_0^\theta a(\mathrm{th}^2\tfrac{1}{2}\theta+\tfrac{1}{4}\mathrm{sech}^4\tfrac{1}{2}\theta)^{\frac{1}{2}}\,d\theta = a\int_0^\theta (1-\tfrac{1}{2}\mathrm{sech}^2\tfrac{1}{2}\theta)\,d\theta = a\theta-r$$

and $$A = \tfrac{1}{2}\int_0^\theta r^2\,d\theta = \tfrac{1}{2}a^2\int_0^\theta d\theta-\tfrac{1}{2}a^2\int_0^\theta \mathrm{sech}^2\tfrac{1}{2}\theta\,d\theta = \tfrac{1}{2}a^2\theta-ar.$$

10.41. $\tan\psi = dy/dx = \tfrac{3}{2}ax^{\frac{1}{2}}$, whence $x = 4\tan^2\psi/9a^2$ and

$$dx/d\psi = 8\tan\psi\sec^2\psi/9a^2,$$

and therefor

$$s = \int_0^x \sec\psi\,dx = \int_0^\psi \sec\psi\frac{dx}{d\psi}\,d\psi = \frac{8}{9a^2}\int_0^\psi \tan\psi\sec^3\psi\,d\psi = \frac{8}{27a^2}\sec^3\psi,$$

and so $27a^2s = 8\sec^3\psi$.

The radius of curvature $= ds/d\psi = (8/9a^2)\sec^3\psi\tan\psi$; furthermore $PG = y/\cos\psi$, $PH = x/\sin\psi$ and so

$$HP^4/PG = (x^4/y)\cos\psi\,\mathrm{cosec}^4\psi = x^{\frac{5}{2}}a^{-1}\,\mathrm{cosec}^4\psi\cos\psi,$$

which is proportional to $\tan^5\psi\cos\psi\,\mathrm{cosec}^4\psi = \sec^3\psi\tan\psi$.

10.42. Let $q = y\cos\psi-x\sin\psi$, then $p = |q|$;

$$dq/d\psi = -y\sin\psi-x\cos\psi+\rho(dy/ds)\cos\psi-\rho(dx/ds)\sin\psi$$
$$= -y\sin\psi-x\cos\psi$$

and so $$q+d^2q/d\psi^2 = -\rho(\sin^2\psi+\cos^2\psi) = -\rho.$$

But $q^2 = p^2$ and so $q\,dq/d\psi = p\,dp/d\psi$, whence $(dq/d\psi)^2 = (dp/d\psi)^2$ and

$$q\,d^2q/d\psi^2+(dq/d\psi)^2 = p\,d^2p/d\psi^2+(dp/d\psi)^2,$$

and therefore $q\,d^2q/d\psi^2 = p\,d^2p/d\psi^2$. It follows that

$$q^2+q\,d^2q/d\psi^2 = p^2+p\,d^2p/d\psi^2, \quad\text{whence}\quad |q+d^2q/d\psi^2| = |p+d^2p/d\psi^2|.$$

If $p = \sin\psi\log(\sec\psi+\tan\psi)-1$, then

$$dp/d\psi = \cos\psi\log(\sec\psi+\tan\psi)+\tan\psi$$

and $p+d^2p/d\psi^2 = \sec^3\psi$, i.e. $ds/d\psi = \rho = \sec^3\psi$, whence $s = \tan\psi$. The reflection in the tangent of the centre of curvature is the point $x+\rho\sin\psi$, $y-\rho\cos\psi$; but $dy/ds = \sin\psi$ and so $dy/d\psi = \sec^3\psi\sin\psi = \tan\psi\sec\psi$, whence $y = \sec\psi+\text{constant}$, and therefore

$$y-\rho\cos\psi = \sec\psi+\text{constant}-\sec\psi = \text{constant}.$$

10.43. (i) Curvature $= d^2y/dx^2/\{1+(dy/dx)^2\}^{\frac{3}{2}} = (x+1/x^3)/\{\frac{1}{2}(x^2+1/x^2)\}^3$

$$= 8x^3/(x^4+1)^2 = 8\kappa, \quad x \geqslant 0,$$

and so $d\kappa/dx = x^2(3-5x^4)/(x^4+1)^3$ whence the point of maximum curvature for $x \geqslant 0$ is $x = \sqrt[4]{\frac{3}{5}}$; when $x < 0$, $\kappa = |x|^3/(x^4+1)^2$ and so $x = -\sqrt[4]{\frac{3}{5}}$ is also a point of maximum curvature. When $x \neq 0$, $\kappa > 0$, and when $x = 0$, $\kappa = 0$, so that $x = 0$ is a point of minimum curvature.

(ii) $y = tx$, $y' = x+tx'$, $y'' = 2x'+tx''$ and so

$$x'y''-x''y' = 2x'^2-xx'' = 2e^{-4t}(1-2t+2t^2).$$

Furthermore $x'^2+y'^2 = e^{-4t}(1-2t+2t^2)$ and so the curvature κ is given by

$$\kappa = 2e^{2t}/(1-2t+2t^2)^{\frac{3}{2}},$$

whence

$$\kappa' = 4e^{2t}(2t^2-6t+3)/(2t^2-2t+1)^{\frac{5}{2}}$$

$$= 8e^{2t}(t-t_1)(t-t_2)/(2t^2-2t+1)^{\frac{5}{2}},$$

where

$$t_1 = \tfrac{1}{2}(3-\sqrt{3}), \qquad t_2 = \tfrac{1}{2}(3+\sqrt{3}).$$

Since $8e^{2t}/(2t^2-2t+1)^{\frac{5}{2}}$ is positive for any t, κ' changes from positive to negative as t increases through t_1 and so the curvature is maximum when

$$t = t_1 = \tfrac{1}{2}(3-\sqrt{3}).$$

10.44. Denote dx/dy by x', d^2x/dy^2 by x'', then

$$2ay = x'(a-x)^2-2xx'(a-x),$$

$$2a = x''(a-x)^2-4x'^2(a-x)-2(a-x)xx''+2xx'^2;$$

at the origin $x = y = 0$, whence $x' = 0$ and $x'' = 2/a$ and so the radius of curvature is $a/2$. Next write $x = a+X$, $y = Y$ then $aY^2 = X^2(a+X)$, i.e. $a(X^2-Y^2)+X^3 = 0$: denote now dY/dX, d^2Y/dX^2,... by Y', Y'',..., then we have in turn

$$2aX-2aYY'+3X^2 = 0, \qquad 2a-2aY'^2-2aYY''+6X = 0,$$

$$-6aY'Y''-2aYY'''+6 = 0,$$

whence, taking $X = Y = 0$, we find $Y'^2 = 1$, $Y'Y'' = 1/a$, i.e. on one branch $Y' = 1$, $Y'' = 1/a$ and on the other $Y' = -1$, $Y'' = -1/a$, giving the same value $(1+1)^{\frac{3}{2}}a = 2^{\frac{3}{2}}a$ for the radius of curvature of each branch.

10.5. Since OQ is perpendicular to PQ, $p = r\sin\phi$ and $t^2 = r^2-p^2$; but

$$p^2+(dp/d\psi)^2 = q^2+(dq/d\psi)^2 = (y\cos\psi-x\sin\psi)^2+(y\sin\psi+x\cos\psi)^2$$

$$= x^2+y^2 = r^2$$

and so $t = |dp/d\psi|$.

10.51. The parabola is $y^2 = 4ax$.

$$PC = \{1+(dy/dx)^2\}^{\frac{3}{2}}/|d^2y/dx^2| = |\sec^3\psi/(d^2y/dx^2)|.$$

But $y\,dy/dx = 2a$ and $y\,d^2y/dx^2+(dy/dx)^2 = 0$, whence

$$\sec^3\psi/(d^2y/dx^2) = -y\sec^3\psi/\tan^2\psi.$$

Further $PG = y/\cos\psi$ and the distance of P from the focus $(a, 0)$ is $\{(x-a)^2+y^2\}^{\frac{1}{2}} = x+a$. Hence $PG/PC = \sin^2\psi$; but

$$\tan^2\psi = (dy/dx)^2 = 4a^2/y^2 = a/x$$

and so $\sin^2\psi = a/(x+a)$.

10.52. We have $x^2+y^2 = a^2(a-3)^2e^{-2at}$,

$$x = -(ax+y), \qquad y = x-ay, \qquad x^2+y^2 = (1+a^2)(x^2+y^2),$$
$$x = (a^2-1)x+2y, \qquad y = -2ax+(a^2-1)y.$$

Hence the total length of the spiral is

$$s = \sqrt{\{(1+a^2)a^2(a-3)^2\}}\int\limits_0^\infty e^{-at}\,dt$$

and so $$s^2 = (1+a^2)(a-3)^2, \quad s > 0, \quad a \neq 3.$$

Hence $$s\frac{ds}{da} = (a-3)(2a-1)(a-1)$$

$$\left(\frac{ds}{da}\right)^2+s\frac{d^2s}{da^2} = (2a-1)(a-1)+(2a-1)(a-3)+2(a-3)(a-1).$$

Accordingly s is minimum at $a = \frac{1}{2}$ and $a = 3$, and maximum at $a = 1$. The curvature κ satisfies

$$\kappa^2 = \frac{(x^2+y^2)^3}{(xy-xy)^2} = (1+a^2)(x^2+y^2) = (1+a^2)a^2(a-3)^2e^{-2at}.$$

Hence at the point $t = 1/a$

$$e^2\kappa^2 = (1+a^2)a^2(a-3)$$

and $$e^2\kappa\frac{d\kappa}{da} = a(a-3)(3a^3-6a^2+2a-3);$$

if $\phi(a) = 3a^3-6a^2+2a-3$ then $\phi(1\frac{1}{2}) = -\frac{27}{8}$, $\phi(2) = 1$ so that $d\kappa/da$ vanishes at some point α between $1\frac{1}{2}$ and 2, and as a increases through α, $d\kappa/da$ changes sign from positive to negative, so that κ is maximum when $a = \alpha$.

10.6. Let $x = X\cos\alpha+Y\sin\alpha+a$, $y = X\sin\alpha-Y\cos\alpha+b$, where α, a, b are constant, then

$$dy/dx = (dy/dX)/(dx/dX) = \{\sin\alpha-(dY/dX)\cos\alpha\}/\{\cos\alpha+(dY/dX)\sin\alpha\}$$

and

$$d^2y/dx^2 = \{d(dy/dx)/dX\}/(dx/dX) = -\{d^2Y/dX^2\}/\{\cos\alpha+(dY/dX)\sin\alpha\}^3,$$

whence

$$(d^2y/dx^2)/\{1+(dy/dx)^2\}^{\frac{3}{2}} = -(d^2Y/dX^2)/\{1+(dY/dX)^2\}^{\frac{3}{2}}.$$

10.61. Since $\dot{x} = -(a\sin\theta\,\dot\theta+c\sin\phi\,\dot\phi)$, $\dot{y} = a\cos\theta\,\dot\theta+c\cos\phi\,\dot\phi$, therefore

$$\begin{aligned} x\dot{y}-\dot{x}y &= a^2\dot\theta+c^2\dot\phi+ac\cos(\theta-\phi)(\dot\theta+\dot\phi) \\ &= a^2\dot\theta+c^2\dot\phi+ac\cos(\theta-\phi)(\dot\phi-\dot\theta)+2ac\dot\theta\cos(\phi-\theta), \\ &= a^2\dot\theta+c^2\dot\phi+ac\cos(\phi-\theta)(\dot\phi-\dot\theta)+2cr\dot\psi-2bc\dot\phi, \end{aligned}$$

since a $-\theta) = r\dot\psi - b\dot\phi$, and so the area enclosed is

$$\tfrac{1}{2}\int_{t_0}^{t_1} (x\dot y - \dot x y)\,dt = \tfrac{1}{2}[a^2\theta + (c^2-2bc)\phi + ac\sin(\phi-\theta) + 2cr\psi]_{t_0}^{t_1} = cr(\psi_1-\psi_0),$$

since θ and ϕ return to their original values.

Thus the area enclosed is proportional to the change in ψ.

This example illustrates the principle of *Amsler's planimeter* The planimeter consists of two rods OA, AC smoothly jointed at A, and of lengths a, c respectively. A wheel of radius r is centred at a point B on AC, such that $AB = b$, and is free to rotate about AB as its axis. The point O is fixed and C describes a closed curve, the wheel rolling on the plane of the curve, neither rod making a complete revolution. Taking any two perpendicular directions through O for axes, if θ and ϕ are the inclinations of OA and AC then the coordinates of C are $x = a\cos\theta + c\cos\phi$, $y = a\sin\theta + c\sin\phi$. When the wheel turns through an angle ψ, the velocity of B is $r\dot\psi$ perpendicular to AC; but the velocity of B relative to A is $b\dot\phi$ and the velocity of A, perpendicular to AB, is $a\dot\theta\cos(\theta-\phi)$, thus

$$r\dot\psi = a\dot\theta\cos(\theta-\phi) + b\dot\phi.$$

Hence the area of the curve described by C is proportional to the angle through which the wheel rolls.

10.7. $dy/dx = \operatorname{sh} x/c$, whence

$$s = \int_0^x \sqrt{(1+\operatorname{sh}^2 x/c)}\,dx = \int_0^x \operatorname{ch} x/c\,dx = c\operatorname{sh} x/c.$$

$$\text{Volume} = \int_{-\xi}^{\xi} \pi y^2\,dx = \pi c^2 \int_{-\xi}^{\xi} \operatorname{ch}^2 x/c\,dx = (\pi c^2/2)\int_{-\xi}^{\xi} (1+\operatorname{ch} 2x/c)\,dx$$

$$= \pi c^2\xi + \pi c^2 \operatorname{sh}(x/c)\operatorname{ch}(x/c) = \pi c(c\xi + s\eta).$$

10.71. Curvature $= e^x/(1+e^{2x})^{\frac{3}{2}} = \kappa$. $d\kappa/dx = e^x(1-2e^{2x})/(1+e^{2x})^{\frac{5}{2}}$, point of maximum curvature is given by $e^{2x} = \frac{1}{2}$, i.e. $x = -\log\sqrt{2}$. Curvature at this point is $2/3\sqrt{3}$.

10.72. Let B, C be the points $(0, c)$, $(0, -c)$ then the perpendicular bisector of BC is the x-axis; if the vertex of the parabolas is the point $(a, 0)$ then the equation of the family of parabolas is $y^2/(x-a) =$ constant. The equation of the family of circles may be put in the form

$$(x^2+y^2-c^2)/x = \text{constant}.$$

On any parabola $2y\dfrac{dy}{dx}\Big/(x-a) = y^2/(x-a)^2$ and on any circle

$$1+2(y/x)\frac{dy}{dx} - y^2/x^2 + c^2/x^2 = 0.$$

At a point where a circle and parabola meet at right angles the slopes m of the circle and m' of the parabola satisfy $mm'+1 = 0$.

Since $m = (y^2-x^2-c^2)/2xy$, $m' = y/2(x-a)$ unless $y = 0$, we have

$$y^2-x^2-c^2+4x(x-a) = 0, \quad \text{i.e.} \quad 3x^2-4ax+y^2 = c^2,$$

which is the equation of an ellipse. Furthermore, $y^2 = 0$ belongs to the

family of parabolas $y^2/(x-a) =$ constant, and $y = 0$ cuts every circle through B, C at right angles.

10.73. If P is the point (x, y), and T is $(x_1, 0)$, G is $(x_2, 0)$ then

$$x_1 - X = -y\,dx/dy, \qquad x_2 - X = y\,dy/dx$$

and so $$TG = |x_2 - x_1| = |y\,dy/dx + y\,dx/dy|.$$

Hence $$y\,dy/dx + y\,dx/dy = \pm ds/d\psi,$$

the ambiguity of sign arising from the fact that it is only the positive values of the two sides which are equal. But $\tan\psi = dy/dx$ and so

$$y\tan\psi + y/\tan\psi = \pm ds/d\psi = \pm(dy/d\psi)(ds/dy),$$

whence $$\pm dy/d\psi = y\sec^2\psi\sin\psi/\tan\psi = y\sec\psi,$$

and so $(1/y)(dy/d\psi) = \pm\sec\psi$, whence $\log y = \pm\log(\sec\psi + \tan\psi) + \log a$, and therefore either $\sec\psi + \tan\psi = y/a$ or $\sec\psi + \tan\psi = a/y$. But

$$\sec\psi + \tan\psi = (1+\sin\psi)/\cos\psi = (1+\tan^2\tfrac{1}{2}\psi + 2\tan\tfrac{1}{2}\psi)/(1-\tan^2\tfrac{1}{2}\psi)$$
$$= (1+\tan\tfrac{1}{2}\psi)/(1-\tan\tfrac{1}{2}\psi),$$

whence either $\tan\frac{1}{2}\psi = (y-a)/(y+a)$ or $\tan\frac{1}{2}\psi = (a-y)/(a+y)$.

From $y/a = \tan\psi + \sec\psi$ we have $(y/a - \tan\psi)^2 = \sec^2\psi = 1 + \tan^2\psi$, and so $2ay\dfrac{dy}{dx} = y^2 - a^2$, i.e. $\{1/(y-a) + 1/(y+a)\}\dfrac{dy}{dx} = 1/a$, whence

$$\log(y^2 - a^2) - x/a = \text{constant} = \log a^2 b \text{ (say)}, \quad \text{i.e. } y^2 = a^2(1 + be^{x/a}).$$

Similarly, if $a/y = \sec\psi + \tan\psi$ then $y^2 = a^2(1 - be^{-x/a})$.

10.74. $$dy/dx = \tan x, \qquad d^2y/dx^2 = \sec^2 x,$$

$$\kappa = |\sec^2 x/(1+\tan^2 x)^{\frac{3}{2}}| = |\cos x|, \ -\tfrac{1}{2}\pi < x < \tfrac{1}{2}\pi, = \cos x.$$

Furthermore

$$s = \int_0^x +\sqrt{(1+\tan^2 x)}\,dx = \int_0^x \sec x\,dx, \ -\tfrac{1}{2}\pi < x < \tfrac{1}{2}\pi, = \log(\sec x + \tan x).$$

10.75. $Y - y = (a-x)dy/dx$ and so

$$dY/ds = dy/ds + (a-x)(d^2y/dx^2)(dx/ds) - dy/ds = (a-x)(dx/ds)(d^2y/dx^2).$$

Thus $(a-x)(dx/ds)(d^2y/dx^2) =$ constant $= c$. Write $dy/dx = p$, then

$$ds/dx = \sqrt{(1+p^2)} \quad \text{and so} \quad dp/dx = d^2y/dx^2 = c\sqrt{(1+p^2)}/(a-x),$$

i.e. $\{1/\sqrt{(1+p^2)}\}\,dp/dx = c/(a-x)$, whence

$$\log\{p + \sqrt{(1+p^2)}\} + c\log(a-x) = \text{constant} = c\log a,$$

since $p = 0$ when $x = 0$; therefore $p + \sqrt{(1+p^2)} = \{a/(a-x)\}^c$ and so $1 + p^2 = \{a/(a-x)\}^{2c} - 2p\{a/(a-x)\}^c + p^2$, i.e. $2p = \{a/(a-x)\}^c - \{(a-x)/a\}^c$, and therefore

$$2y = a^c/(c-1)(a-x)^{c-1} + (a-x)^{c+1}/(c+1)a^c - 2ac/(c^2-1)$$

provided $c \neq 1$; if $c = 1$,

$$2y = -a\log(a-x) + (a-x)^2/2a + a\log a - \tfrac{1}{2}a.$$

10.76. Take the tangent to the circle at O, and OA for x- and y-axis respectively, then the circle is $x^2+y^2 = 2ay$ and the tangent at A is $y = 2a$, a being the radius of the circle. Let (r,θ) be the polar coordinates of P, $0 < \theta < \pi$, then $r = QT = OT-OQ = 2a/\sin\theta - 2a\sin\theta$, and so the cartesian coordinates of P are $x = 2a\cot\theta - 2a\sin\theta\cos\theta$, $y = 2a-2a\sin^2\theta$; hence the area between the curve on which P lies and the tangent at A is

$$\left|\int (2a-y)\,dx\right| = \left|\int (2a-y)\frac{dx}{d\theta}\,d\theta\right| = \int_0^\pi 2a\sin^2\theta\{2a\operatorname{cosec}^2\theta + 2a\cos 2\theta\}\,d\theta$$

$$= 4a^2\pi + 2a^2\int_0^\pi (1-\cos 2\theta)\cos 2\theta\,d\theta$$

$$= 4a^2\pi - a^2\int_0^\pi (1+\cos 4\theta)\,d\theta = 3\pi a^2.$$

10.77. At $x = h$, $dy/dx = a_0$, $d^2y/dx^2 = a_1$.

10.78. $$y^2 = x^2(x-2)/(x-4) = x^2+2x+8+32/(x-4)$$
therefore,

$$y\,dy/dx = x+1-16/(x-4)^2, \qquad (dy/dx)^2+y\,d^2y/dx^2 = 1+32/(x-4)^3,$$
$$3(dy/dx)(d^2y/dx^2)+y\,d^3y/dx^3 = -96/(x-4)^4,$$

whence at $x = 0$, $y = 0$ from the last two equations $(dy/dx)^2 = \frac{1}{2}$, $d^2y/dx^2 = \mp\sqrt{2}/8$ and so the curvature is $(\sqrt{2}/8)/(1+\frac{1}{2})^{\frac{3}{2}} = 1/2.3^{\frac{3}{2}} = 1/6\sqrt{3}$.

10.79. Any circle through $(-1,0)$, $(3,0)$ has the equation

$$(x-1)^2+(y-c)^2 = c^2+4, \quad \text{i.e.} \quad \{(x-1)^2+y^2-4\}/y = \text{constant}.$$

Normal at (x,y) to circle is $(X-x)+(Y-y)\dfrac{dy}{dx} = 0$ which passes through the origin $X = Y = 0$ if $x+y\dfrac{dy}{dx} = 0$; but on the circle

$$2(x-1)/y-\{(x-1)^2/y^2\}\,dy/dx+dy/dx+(4/y^2)\,dy/dx = 0,$$

whence writing $dy/dx = -x/y$, we have $y^2(x-2)+x(x+1)(x-3) = 0$.

10.791. $2(x^2+y^2)(x+y\,dy/dx) = a^2(y\,dy/dx-x)$ and differentiating again,

$$4(x+y\,dy/dx)^2+2(x^2+y^2)[1+(dy/dx)^2+y\,d^2y/dx^2]$$
$$= a^2[(dy/dx)^2+y\,d^2y/dx^2-1],$$

whence taking $x = 0$, $y = a$ we have $2a^3\,dy/dx = a^3\,dy/dx$, so that $dy/dx = 0$ and therefore

$$2a^2(1+a\,d^2y/dx^2) = a^2(a\,d^2y/dx^2-1),$$

i.e. $$a\,d^2y/dx^2 = -3,$$

so the curvature is $3/a$.

10.792. For the cycloid $x = a(t+\sin t)$, $y = a(1-\cos t)$ we have

$$dy/dx = \sin t/(1+\cos t) = \tan\tfrac{1}{2}t$$

and $$d^2y/dx^2 = \tfrac{1}{2}\sec^2\tfrac{1}{2}t/a(1+\cos t) = 1/4a\cos^4\tfrac{1}{2}t.$$

Hence the centre of curvature (X,Y) is given by

$$X = a(t+\sin t)-4a\sec^2\tfrac{1}{2}t\tan\tfrac{1}{2}t\cos^4\tfrac{1}{2}t = a(t+\sin t)-2a\sin t = a(t-\sin t)$$
$$Y = a(1-\cos t)+4a\sec^2\tfrac{1}{2}t\cos^4\tfrac{1}{2}t = 2a+a(1+\cos t)$$

so that (X,Y) lies on an equal cycloid.

10.793. On any curve of the family, $3x^2+2y\,dy/dx = 0$; on the circle $x^2+y^2 = a$, $x+y\,dy/dx = 0$. Thus a circle and a curve of the family intersect at right angles at a point (x,y) satisfying $(-3x^2/2y)(-x/y)+1 = 0$, i.e. $3x^3+2y^2 = 0$. For contact we require $3x^2 = 2x$, i.e. either $x = 0$ or $x = \frac{2}{3}$.

Observe that one, and only one, curve of each family passes through *any* point; for if x and y are given then c and a are determined uniquely by the equations $c = x^3+y^2$, $a = x^2+y^2$.

10.794. $\dot{s} = +\sqrt{(\dot{x}^2+\dot{y}^2)} = \sqrt{\{(2-\sec^2 t)^2+4\tan^2 t\}} = \sec^2 t$

and so $s = \tan t$. Furthermore

$$\tan\psi = \dot{y}/\dot{x} = 2\tan t/(2-\sec^2 t) = 2\tan t/(1-\tan^2 t) = \tan 2t$$

whence $\sec^2\psi\,.\,\dot{\psi} = 2\sec^2 2t$, and so $\dot{\psi} = 2$.

The radius of curvature $= \left|\dfrac{ds}{d\psi}\right| = \dot{s}/\dot{\psi} = \frac{1}{2}\dot{s} = \frac{1}{2}(1+s^2)$.

10.795. $2\dfrac{dy}{dx} = x^3-1/x^3$, therefore

$$2\frac{ds}{dx} = \sqrt{\left\{4+4\left(\frac{dy}{dx}\right)^2\right\}} = \sqrt{(x^3+1/x^3)^2} = x^3+1/x^3, \quad x > 0,$$

and so the required surface area

$$= 2\pi\int_1^2 x\frac{ds}{dx}\,dx = \pi\int_1^2 x(x^3+1/x^3)\,dx = 6{\cdot}7\pi.$$

10.8. On the cycloid $s = 4a\sin\psi$, and so the centroid is given by

$$\bar{x} = \int_0^{\frac{1}{2}\pi} x\frac{ds}{d\psi}\,d\psi\Big/4a = a\int_0^{\frac{1}{2}\pi}(2\psi+\sin 2\psi)\cos\psi\,d\psi = a(\pi-2+\tfrac{2}{3}) = (\pi-\tfrac{4}{3})a,$$

$$\bar{y} = \int_0^{\frac{1}{2}\pi} y\frac{ds}{d\psi}\,d\psi\Big/4a = a\int_0^{\frac{1}{2}\pi}(1-\cos 2\psi)\cos\psi\,d\psi = 2a\int_0^{\frac{1}{2}\pi}\sin^2\psi\cos\psi\,d\psi = 2a/3.$$

10.81. On $y = \operatorname{ch} x$, $dy/dx = \operatorname{sh} x$ and $s = \operatorname{sh} x$ and therefore

$$\bar{x} = \int_0^x x(k/y^2)\frac{ds}{dx}\,dx\Big/\int_0^x (k/y^2)\frac{ds}{dx}\,dx = \int_0^x x\operatorname{sech}^2 x\,dx\Big/\int_0^x \operatorname{sech}^2 x\,dx$$

$$= x-\log\operatorname{ch} x/\operatorname{th} x = x-(y/s)\log y.$$

and

$$\bar{y} = \int_0^x y(k/y^2)\frac{ds}{dx}\,dx\Big/\int_0^x (k/y^2)\frac{ds}{dx}\,dx = \int_0^x (1/\operatorname{ch} x)\,dx\Big/\operatorname{th} x$$

$$= \int_0^x \{\operatorname{ch} x/(1+\operatorname{sh}^2 x)\}\,dx\Big/\operatorname{th} x = (y/s)\tan^{-1}(\operatorname{sh} x) = (y/s)\tan^{-1} s.$$

When $x = n$, $y/s = (e^n+e^{-n})/(e^n-e^{-n}) = (1+e^{-2n})/(1-e^{-2n}) \to 1$ and $\tan^{-1}\operatorname{sh} n = \sin^{-1}\operatorname{th} n \to \frac{1}{2}\pi$ since $\operatorname{th} n \to 1$, so that $\bar{y} \to \frac{1}{2}\pi$; also

$$\bar{x} = n-\{(e^n-e^{-n})/(e^n+e^{-n})\}\log\{(e^n+e^{-n})/2\}$$
$$= n-\{(e^n-e^{-n})/(e^n+e^{-n})\}\{n-\log 2+\log(1+e^{-2n})\}$$
$$= 2ne^{-n}/(e^n+e^{-n})+\operatorname{th} n\{\log 2-\log(1+e^{-2n})\}$$
$$= (2n/e^{2n})/(1+e^{-2n})+\operatorname{th} n\{\log 2-\log(1+e^{-2n})\} \to \log 2.$$

10.811. $$\bar{x} = \int_0^{\frac{1}{2}\pi} xy\,dx \Big/ \int_0^{\frac{1}{2}\pi} y\,dx = \int_0^{\frac{1}{2}\pi} x\sin x\,dx \Big/ \int_0^{\frac{1}{2}\pi} \sin x\,dx = 1,$$

$$\bar{y} = \tfrac{1}{2}\int_0^{\frac{1}{2}\pi} y^2\,dx \Big/ \int_0^{\frac{1}{2}\pi} y\,dx = \tfrac{1}{2}(\tfrac{1}{4}\pi-1) = \tfrac{1}{8}\pi-\tfrac{1}{2}.$$

10.812. Let the fixed tangent be the x-axis, and the equation of the circumference $x^2+y^2 = 2ax$, i.e. the pair of curves $y = y_1 = +\sqrt{(2ax-x^2)}$, $y = y_2 = -\sqrt{(2ax-x^2)}$. Then $\bar{y} = 0$, since $y_1^2 = y_2^2$ and

$$\bar{x} = \int_0^{2a} 2x\sqrt{(2ax-x^2)}kx\,dx \Big/ \int_0^{2a} 2\sqrt{(2ax-x^2)}kx\,dx$$

$$= 2a^4\int_0^{\frac{1}{2}\pi} 32\sin^6\theta\cos^2\theta\,d\theta \Big/ 2a^3\int_0^{\frac{1}{2}\pi} 16\sin^4\theta\cos^2\theta\,d\theta, \quad x = 2a\sin^2\theta,$$

$$= 5a/4$$

and so the centroid is distant $5a/4-a = \frac{1}{4}a$ from the centre.

10.82. $$\int_a^b 2\pi y\,ds = \left\{2\pi\left(\int_a^b y\,ds \Big/ \int_a^b ds\right)\right\}\left(\int_a^b ds\right).$$

10.83. $$\int_a^b \pi y^2\,dx = \left\{2\pi\left(\int_a^b \tfrac{1}{2}y^2\,dx \Big/ \int_a^b y\,dx\right)\right\}\left(\int_a^b y\,dx\right).$$

10.84. In the parabola $y^2 = 4ax$, $ds/dx = \{1+(dy/dx)^2\}^{\frac{1}{2}} = 2\{a(a+x)\}^{\frac{1}{2}}/y$ and so

$$\bar{x} = \int_0^{3a} 2\pi yx\frac{ds}{dx}\,dx \Big/ \int_0^{3a} 2\pi y\frac{ds}{dx}\,dx$$

$$= \int_0^{3a} x\{a(a+x)\}^{\frac{1}{2}}\,dx \Big/ \int_0^{3a} \{a(a+x)\}^{\frac{1}{2}}\,dx$$

$$= 3\int_{\sqrt{a}}^{2\sqrt{a}} 2(t^2-a)t^2\,dt \Big/ 14a^{\frac{3}{2}}, \quad t^2 = a+x,$$

$$= [t^3(3t^2-5)]_{\sqrt{a}}^{2\sqrt{a}} \Big/ 35a^{\frac{3}{2}} = 58\ldots$$

10.9. If $f''(x) > 0$ the radius of curvature has the value

$$+\{1+[f'(x)]^2\}^{\frac{3}{2}}/f''(x)$$

and if $f''(x) < 0$ its value is $-\{1+[f'(x)]^2\}^{\frac{3}{2}}/f''(x)$; but σ has the same sign as $f''(x)$ and $|\sigma|$ equals the radius of curvature, and therefore

$$\sigma = +\{1+[f'(x)]^2\}^{\frac{3}{2}}/f''(x).$$

Furthermore
$$\cos\psi = 1/+\{1+[f'(x)]^2\}^{\frac{1}{2}}$$
and so
$$\begin{aligned} a &= x-f'(x)\{1+[f'(x)]^2\}/f''(x) \\ &= x-f'(x)\{1+[f'(x)]^2\}^{\frac{3}{2}}/f''(x)\{1+[f'(x)]^2\}^{\frac{1}{2}}, \end{aligned}$$

both square roots being positive,

$$= x-\sigma\tan\psi\cos\psi = x-\sigma\sin\psi.$$

Similarly
$$\begin{aligned} b &= y+\{1+[f'(x)]^2\}/f''(x) \\ &= y+\{1+[f'(x)]^2\}^{\frac{3}{2}}/f''(x)\{1+[f'(x)]^2\}^{\frac{1}{2}} \\ &= y+\sigma\cos\psi. \end{aligned}$$

10.91.
$$x = a\cos^3 t, \qquad y = a\sin^3 t$$
and so
$$x' = -3a\cos^2 t\sin t, \qquad y' = 3a\sin^2 t\cos t.$$
Length of a quadrant

$$= \int_0^{\frac{1}{2}\pi} (x'^2+y'^2)^{\frac{1}{2}}\,dt = 3a\int_0^{\frac{1}{2}\pi} \sin t\cos t\,dt = \frac{3a}{2}\int_0^{\frac{1}{2}\pi} \sin 2t\,dt = \frac{3a}{2}.$$

Radius of curvature

$$\begin{aligned} &= \{1+(dy/dx)^2\}^{\frac{3}{2}}/(d^2y/dx^2) = x'\{1+(y'/x')^2\}^{\frac{3}{2}}/D_t(y'/x') \\ &= \sec^3 t\,.\,3a\cos^2 t\sin t/\sec^2 t = \tfrac{3}{2}a\sin 2t; \end{aligned}$$

hence maximum radius of curvature is $\frac{3}{2}a$.

Area of the surface of revolution

$$= 2\pi\int y\,ds = 2\pi\int_0^{\pi} ys'\,dt = 4\pi\int_0^{\frac{1}{2}\pi} ys'\,dt = 12\pi a^2\int_0^{\frac{1}{2}\pi} \sin^4 t\cos t\,dt = 12\pi a^2/5.$$

10.92. The transformation is given by

$$X = x\cos\alpha+y\sin\alpha+k, \qquad Y = x\sin\alpha-y\cos\alpha+l.$$

Then

$$\begin{aligned} &\tfrac{1}{2}\int_{t_0}^{t_1} (XY'-X'Y)\,dt \\ &= \tfrac{1}{2}\int_{t_0}^{t_1} \{(x\cos\alpha+y\sin\alpha+k)(x'\sin\alpha-y'\cos\alpha)- \\ &\qquad\qquad -(x\sin\alpha-y\cos\alpha+l)(x'\cos\alpha+y'\sin\alpha)\}\,dt \\ &= \tfrac{1}{2}\int_{t_0}^{t_1} (x'y-xy')\,dt+\tfrac{1}{2}(k\sin\alpha-l\cos\alpha)\int_{t_0}^{t_1} x'\,dt-\tfrac{1}{2}(k\cos\alpha+l\sin\alpha)\int_{t_0}^{t_1} y'\,dt \\ &= \tfrac{1}{2}\int_{t_0}^{t_1} (x'y-xy')\,dt \quad \text{for } x(t_1) = x(t_0),\ y(t_1) = y(t_0), \end{aligned}$$

since the curve is closed.

10.93. The evolute is

$$x = a\mu^2+(4a^2\mu^2+4a^2)2a/4a^2 = 3a\mu^2+2a,$$
$$y = 2a\mu-(4a^2\mu^2+4a^2)2a\mu/4a^2 = -2a\mu^3,$$

i.e. $4(x-2a)^3 = 27ay^2$.

10.94. The evolute of the curve $x = x(s)$, $y = y(s)$ is

$$X = x-y'(x'^2+y'^2)/(x'y''-x''y'), \qquad Y = y+x'(x'^2+y'^2)/(x'y''-x''y').$$

But $x' = \dfrac{dx}{ds} = \cos\psi$, $y' = \sin\psi$ and so $x'' = -\sin\psi\dfrac{d\psi}{ds}$, $y'' = \cos\psi\dfrac{d\psi}{ds}$ so that $X = x-\sin\psi\dfrac{ds}{d\psi}$, $Y = y+\cos\psi\dfrac{ds}{d\psi}$. and therefore

$$\left(\frac{dS}{ds}\right)^2 = \left(\frac{dX}{ds}\right)^2+\left(\frac{dY}{ds}\right)^2 = \left(\frac{d^2s}{d\psi^2}\frac{d\psi}{ds}\right)^2,$$

$$\frac{dX}{ds} = -\frac{d^2s}{d\psi^2}\frac{d\psi}{ds}\sin\psi, \qquad \frac{dY}{ds} = \frac{d^2s}{d\psi^2}\frac{d\psi}{ds}\cos\psi$$

Hence $\dfrac{dX}{dS} = \mp\sin\psi$, $\dfrac{dY}{dS} = \pm\cos\psi$, that sign being chosen which make $\dfrac{dX}{dS}$ positive. Thus $\cos\Psi\cos\psi+\sin\Psi\sin\psi = 0$, i.e. $\cos(\Psi-\psi) = 0$; but $-\pi \leqslant \Psi-\psi \leqslant \pi$ and so $\Psi = \psi\pm\frac{1}{2}\pi$, choosing that sign which leaves Ψ in $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, i.e. $\Psi = \psi+\frac{1}{2}\pi$ if $\sin\psi < 0$ and $\Psi = \psi-\frac{1}{2}\pi$ if $\sin\psi \geqslant 0$. Furthermore

$$S = \int\frac{dS}{ds}\,ds = \pm\int\frac{d^2s}{d\psi^2}\frac{d\psi}{ds}\,ds = \pm\int\frac{d^2s}{d\psi^2}\,d\psi = \pm\frac{ds}{d\psi},$$

choosing the upper sign if $ds/d\psi$ is positive, and the lower if $ds/d\psi$ is negative.

10.95. Cartesian coordinates of P relative to O = coordinates of N relative to O+coordinates of P relative to N, and so

$$x_P = c\,\mathrm{sh}\,\psi\cos\psi-c\,\mathrm{ch}\,\psi\sin\psi, \qquad y_P = c\,\mathrm{sh}\,\psi\sin\psi+c\,\mathrm{ch}\,\psi\cos\psi,$$

and similarly

$$x_Q = c\,\mathrm{sh}\,\psi\cos\psi+c\,\mathrm{ch}\,\psi\sin\psi, \qquad y_Q = c\,\mathrm{sh}\,\psi\sin\psi-c\,\mathrm{ch}\,\psi\cos\psi.$$

Denoting differentiation with respect to ψ by a dot, we have

$$\dot{x}_Q = 2c\,\mathrm{ch}\,\psi\cos\psi, \quad \dot{y}_Q = 2c\,\mathrm{ch}\,\psi\sin\psi, \quad \ddot{x}_Q = 2c(\mathrm{sh}\,\psi\cos\psi-\mathrm{ch}\,\psi\sin\psi),$$

and $$\ddot{y}_Q = 2c(\mathrm{ch}\,\psi\cos\psi+\mathrm{sh}\,\psi\sin\psi).$$

Hence $$\dot{x}_Q\ddot{y}_Q-\ddot{x}_Q\dot{y}_Q = 4c^2\,\mathrm{ch}^2\psi = \dot{x}_Q^2+\dot{y}_Q^2,$$

and the evolute is

$$X = x_Q-\dot{y}_Q = x_P, \qquad Y = y_Q+\dot{x}_Q = y_P.$$

Furthermore $s = 2c\int_0^\psi \mathrm{ch}\,\psi\,d\psi = 2c\,\mathrm{sh}\,\psi$, and

$$r^2-c^2 = c^2\mathrm{sh}^2\psi+c^2\mathrm{ch}^2\psi-c^2 = 2c^2\mathrm{sh}^2\psi = \tfrac{1}{2}s^2.$$

10.96. $\dot{x} = -(x+y)$, $\dot{y} = x-y$ and $\ddot{x} = 2y$, $\ddot{y} = -2x$, therefore

$$\dot{x}^2+\dot{y}^2 = 2(x^2+y^2) = 4e^{-2t},$$

and so
$$s = 2\int_0^\infty e^{-2t}\,dt = 2[-e^{-t}]_0^\infty = 2.$$

Furthermore

$$\dot{x}\ddot{y}-\ddot{x}\dot{y} = 2x(x+y)-2y(x-y) = 2(x^2+y^2) = \dot{x}^2+\dot{y}^2$$

and therefore the centre of curvature (X, Y) is given by

$$X = x-\dot{y} = y, \qquad Y = y+\dot{x} = -x.$$

Hence the locus of (X, Y) is the spiral on which (x, y) lies rotated about the origin clockwise through a right angle.

10.97. Let the vertices be numbered in order from 1 to n. Initially the side $n1$ rests on the line, then the side 12, then 23, and so on. The polygon turns about the vertices 1, 2, 3,... in turn, turning through an angle $2\pi/n$ about each vertex. When the polygon turns about the vertex r, the vertex n describes an arc of a circle of angle $2\pi/n$ and of radius rn. The diagonal rn subtends an angle $2r\pi/n$ at the centre, and so the length of the diagonal is $2a\sin(r\pi/n)$. Hence the length of the path of the vertex n, in one revolution of the polygon, is $4a\sum^{n-1}\frac{\pi}{n}\sin\frac{r\pi}{n}$.

But

$$\sum_1^{n-1}\frac{\pi}{n}\sin\frac{r\pi}{n} \to \int_0^\pi \sin x\,dx, \quad \text{by Theorem 9.141,}$$

$$= [-\cos x]_0^\pi = 2.$$

Thus the length of an arch of the cycloid, the locus of a point on the circumference of a rolling circle of radius a, is $8a$.

The area bounded by the path of the vertex n, in one revolution, and the line on which the polygon rolls, is the sum of the areas of the $n-1$ sectors with centres 1, 2, 3,..., $n-1$, angle $2\pi/n$, and radii $2a\sin(r\pi/n)$, $r = 1, 2, 3,\ldots, n-1$, together with the sum of the areas of the triangles $12n$, $23n$, $34n$,..., $(n-2)(n-1)n$, i.e.

$$\sum\frac{1}{2}\cdot\frac{2\pi}{n}\cdot 4a^2\sin^2\frac{r\pi}{n}+\frac{n}{2}a^2\sin\frac{2\pi}{n} = 4a^2\sum\frac{\pi}{n}\sin^2\frac{r\pi}{n}+\frac{n}{2}a^2\sin\frac{2\pi}{n},$$

but

$$\sum\frac{\pi}{n}\sin^2\frac{r\pi}{n} \to \int_0^\pi \sin^2 x\,dx = \tfrac{1}{2}\pi \quad\text{and}\quad \frac{na^2}{2}\sin\frac{2\pi}{n} = \pi a^2\frac{\sin 2\pi/n}{2\pi/n} \to \pi a^2.$$

and so the area bounded by an arch of the cycloid and the base line is $2\pi a^2+\pi a^2 = 3\pi a^2$.

10.98. Let (α, β) be a current point on S^*, and denote differentiation with respect to the arc length $\sigma = \sigma(t)$ of S^* by a dot, and differentiation with respect to the parameter t by a dash.

Any point (X,Y) on the tangent to S^* at (α,β) satisfies

$$X = \alpha+q\dot\alpha, \qquad Y = \beta+q\dot\beta,$$

where $|q|$ is the distance from (α,β) to (X,Y).

Let the points P_1, P_2 correspond to $q = q_1$, $q = q_2$ respectively, and le $X_1 = \alpha+q_1\dot\alpha$, etc.

The normal to S^* at (α,β) is $(x-\alpha)\dot\alpha+(y-\beta)\dot\beta = 0$ and the normal t S at (X,Y) is $(x-X)X'+(y-Y)Y' = 0$, and so the normals at P_1, P_2 and (

$$\begin{vmatrix} X_1' & Y_1' & X_1X_1'+Y_1Y \\ X_2' & Y_2' & X_2X_2'+Y_2Y \\ \dot\alpha & \dot\beta & \alpha\dot\alpha+\beta\dot\beta \end{vmatrix}$$

Now $$XX'+YY' = \alpha X'+\beta Y'+q(\dot\alpha X'+\dot\beta Y'),$$

and $$\dot\alpha X+\dot\beta Y = \alpha\dot\alpha+\beta\dot\beta+q, \quad \text{since } \dot\alpha^2+\dot\beta^2 = 1,$$

and therefore

$$(\dot\alpha X'+\dot\beta Y')\frac{dt}{d\sigma} = \alpha\ddot\alpha+\beta\ddot\beta+1+\dot q-\ddot\alpha X-\ddot\beta Y = 1+\dot q \quad \text{since } \dot\alpha\ddot\alpha+\dot\beta\ddot\beta = 0.$$

In the determinant Δ, taking α times the first column and β times tl second om tne tnira colum

$$\frac{\sigma}{t}\cdot\begin{vmatrix} X_1' & Y_1' & q_1(1+ \\ X_2' & Y_2' & q_2(1+ \\ \dot\alpha & \dot\beta & 0 \end{vmatrix} = \begin{vmatrix} X_1' & Y_1' & q_1(1+ \\ X_2' & Y_2' & q_2(1+ \\ \alpha' & \beta' & 0 \end{vmatrix}$$

$$\alpha'Y' = D_s(\beta'X-\alpha'Y)-(\beta''X-\alpha''Y)$$

$$= D_s(\alpha\beta'-\alpha'\beta)-(\alpha\beta''-\alpha''\beta)-q(\alpha'\beta''-\alpha''\beta')$$

$$= -q(\alpha'\beta''-\alpha''\beta') \quad d\sigma$$

refore

$$\Delta = \{-q_1q_2(1+\dot q_1)(\alpha'\beta''-\alpha''\beta')+q_1q_2(1+\dot q_2)(\alpha'\beta''-\alpha''\beta')$$

$$= q_1q_2(\dot q_2-\dot q_1)(\alpha'\beta''-\alpha''\beta')\frac{dt}{d\sigma} = 0,$$

for q_2-q_1 is either maximum or minimum and so $\dot q_2-\dot q_1 = 0$.

XI

11.1. (i) $(D-1)^2ye^{-x} = x+x\cos x$; consider $(D-1)^2z = x$, $z = ye^{-x}$, then $z = (1+2D)x = x+2$.

Consider next $(D-1)^2z = \sin ax$;

$$(D-1)^2(D+1)^2\sin ax = (D^2-1)^2\sin ax = (1+a^2)^2\sin ax$$

and

$$(D+1)^2\sin ax = (D^2+1)\sin ax+2D\sin ax = (1-a^2)\sin ax+2a\cos ax.$$

Thus a solution of $(D-1)^2z = \sin ax$ is

$$z = \{(1-a^2)\sin ax+2a\cos ax\}(1+a^2)^{-2};$$

differentiating with respect to a we find that a solution

$$(D-1)^2u = x\cos ax$$

is

$$u = dz/da = -2a(1+a^2)^{-2}\sin ax - 4a(1-a^2)(1+a^2)^{-3}\sin ax -$$
$$-2xa(1+a^2)^{-2}\sin ax + x(1-a^2)(1+a^2)^{-2}\cos ax + 2(1+a^2)^{-2}\cos ax -$$
$$-8a^2(1+a^2)^{-3}\cos ax$$

and so a solution of $(D-1)^2ye^{-x} = x\cos x$ is $ye^{-x} = -\frac{1}{2}[(1+x)\sin x+\cos x]$. Thus the general solution of (i) is

$$y = (A+Bx)e^{2x}+e^{x}(x+2)-\tfrac{1}{2}e^{x}(\sin x+\cos x+x\sin x).$$

(ii) Since $(D^4+D^2+1)\cos ax = (1-a^2+a^4)\cos ax$, therefore a solution of $(D^4+D^2+1)z = \cos ax$ is $z = \cos ax/(1-a^2+a^4)$, whence a solution of $(D^4+D^2+1)u = -x\sin ax$ is

$$u = dz/da = -[x\sin ax(1-a^2+a^4)+(4a^3-2a)\cos ax]/(1-a^2+a^4)^2$$

and so a particular solution of (ii) is $y = x\sin x+2\cos x$.

Since

$$t^4+t^2+1 = (t^2+t+1)(t^2-t+1) = \{(t+\tfrac{1}{2})^2+(\sqrt{3}/2)^2\}\{(t-\tfrac{1}{2})^2+(\sqrt{3}/2)^2\}$$

it follows that the general solution is

$$y = e^{\frac{1}{2}x}\{A\cos\tfrac{1}{2}\sqrt{3}\,x+B\sin\tfrac{1}{2}\sqrt{3}\,x\}+e^{-\frac{1}{2}x}\{C\cos\tfrac{1}{2}\sqrt{3}\,x+E\sin\tfrac{1}{2}\sqrt{3}\,x\}$$
$$+x\sin x+2\cos x.$$

(iii) Transposing e^{3x} we have $(D^2+4)^2ye^{-3x} = \sin x+\sin 2x$; a solution of $(D^2+4)^2z = \sin x$ is $z = \sin x/9$, and a solution of $(D^2+4)^2z = \sin 2x$ is $z = x^2\sin(2x+3\pi)/4^2.2! = -x^2\sin 2x/32$, hence the general solution is

$$y = e^{3x}\{(L+Mx)\sin 2x+(P+Qx)\cos 2x+\tfrac{1}{9}\sin x-\tfrac{1}{32}x^2\sin 2x\}$$
$$= e^{3x}\{(L+Mx-\tfrac{1}{32}x^2)\sin 2x+(P+Qx)\cos 2x+\tfrac{1}{9}\sin x\}.$$

11.2. $dz/dx = 1+dy/dx$ and so $x\,dz/dx = z\log z$ or $(1/z\log z)\,dz/dx = 1/x$, whence $\log\log z = \log x+\log c$, i.e. $\log z = cx$, and so $x+y = e^{cx}$.

11.21.
$$dz/dx = x\,d^2y/dx^2+dy/dx-dy/dx = x\,d^2y/dx^2$$
$$x\,dz/dx = (2x^2+3)z, \quad \text{i.e.} \quad (1/z)dz/dx = 2x+3/x,$$

and so

$$\log z = x^2+3\log x+\log 2a, \quad \text{whence} \quad x\,dy/dx-y = z = 2ax^3e^{x^2},$$

i.e. $dy/dx-y/x = 2ax^2e^{x^2}$; the integrating factor is $e^{-\int dx/x} = e^{-\log x} = 1/x$, therefore $y/x = a\int 2xe^{x^2}\,dx = ae^{x^2}+b$, and so $y = axe^{x^2}+bx$.

11.3.
$$dy/dx = (dy/dt)(dt/dx) = (1/x)dy/dt,$$
$$d^2y/dx^2 = (-1/x^2)dy/dt+(1/x^2)d^2y/dt^2,$$

i.e. $x\,dy/dx = dy/dt$, $x^2\,d^2y/dx^2 = d^2y/dt^2-dy/dt$; in fact if D^ny denotes d^ny/dx^n and Δ^ny denotes d^ny/dt^n then $x^nD^ny = \Delta(\Delta-1)(\Delta-2)\ldots(\Delta-n+1)y$, for if this is true for $n = p$, then

$$D^{p+1}y = D\{(1/x^p)\Delta(\Delta-1)\ldots(\Delta-p+1)y\}$$
$$= -(p/x^{p+1})\Delta(\Delta-1)\ldots(\Delta-p+1)y+(1/x^{p+1})\Delta^2(\Delta-1)\ldots(\Delta-p+1)y$$
$$= (1/x^{p+1})\Delta(\Delta-1)\ldots(\Delta-p+1)(\Delta-p)y.$$

Hence $(x^2D^2-xD+1)y = (\Delta^2-\Delta-\Delta+1)y = (\Delta-1)^2y$, and the equation becomes $(\Delta-1)^2y = e^t$, i.e. $\Delta^2ye^{-t} = 1$, whence

$$y = e^t(A+Bt+\tfrac{1}{2}t^2) = x\{A+B\log x+\tfrac{1}{2}(\log x)^2\}.$$

11.31. $$dz/dx = -ye^{-\int y\,dx} = -yz,$$

$$d^2z/dx^2 = -z\,dy/dx-y\,dz/dx = -z(dy/dx-y^2),$$

and so $d^2z/dx^2+z(2+3y) = 0$, whence $d^2z/dx^2-3dz/dx+2z = 0$, i.e. $(D^2-3D+2)z = 0$, and so $z = Ae^x+Be^{2x}$; therefore

$$yz = -dz/dx = -Ae^x-2Be^{2x}$$

and so

$$y = -(Ae^x+2Be^{2x})/(Ae^x+Be^{2x}) = -(2e^x+c)/(e^x+c) = -2+c/(e^x+c),$$
$c = A/B$.

11.4. Write $dy/dx = p$, then $d^2y/dx^2 = dp/dx$, and so $(dp/dx)^2 = 4p$, whence *either* $dp/dx = -2p^{\frac{1}{2}}$ *or* $dp/dx = 2p^{\frac{1}{2}}$; if the former $p^{\frac{1}{2}} = a-x$ and if the latter $p^{\frac{1}{2}} = a+x$; but $p = 0$ when $x = 0$ and so $a = 0$ and $p = x^2$ in either case, from which it follows that $y = \frac{1}{3}x^3+1$, since $y = 1$ when $x = 0$.

11.41. Write $dy/dx = p$ then $d^2y/dx^2 = (dp/dy)dy/dx = p\,dp/dy$ and the equation becomes $p\,dp/dy = 2\sin 2y$, whence $p^2 = -2\cos 2y+a$; but $p = 2$ and $y = \frac{1}{2}\pi$ when $x = 0$ so that $4 = 2+a$, i.e. $a = 2$, and

$$p^2 = 2(1-\cos 2y) = 4\sin^2 y,$$

whence $p = 2\sin y$, the positive root being chosen since $p = +2$ when $y = \frac{1}{2}\pi$. Hence $\operatorname{cosec} y\,dy/dx = 2$, $0 < y < \pi$, whence

$$\log\tan\tfrac{1}{2}y = 2x+a = 2x$$

since $y = \frac{1}{2}\pi$ when $x = 0$; therefore $\tan\frac{1}{2}y = e^{2x}$ and so

$$\sin y = 2e^{2x}/(1+e^{4x}) = 2/(e^{2x}+e^{-2x}) = \operatorname{sech} 2x,$$

and

$$\cos y = (1-e^{4x})/(1+e^{4x}) = -(e^{2x}-e^{-2x})/(e^{2x}+e^{-2x}) = -\operatorname{sh} x/\operatorname{ch} x.$$

Hence when $0 < y < \frac{1}{2}\pi, x < 0$, and when $\frac{1}{2}\pi < y < \pi, x > 0$; accordingly $y = \cos^{-1}(-\operatorname{th} 2x) = \pi-\cos^{-1}(\operatorname{th} 2x)$. The solution $y = \sin^{-1}(\operatorname{sech} 2x)$ is valid only for $x \leqslant 0$, and when $x > 0$, $y = \pi-\sin^{-1}(\operatorname{sech} 2x)$.

11.5. If $y = e^{x^n}$, $dy/dx = nx^{n-1}y$, $d^2y/dx^2 = n(n-1)x^{n-2}y+n^2x^{2n-2}y$, and so $2+4x^2 = n(n-1)x^{n-2}+n^2x^{2n-2}$, which is satisfied by $n = 2$, so that $y = e^{x^2}$ is a solution. Write $y = ue^{x^2}$, where u is a function of x; then $Dy = e^{x^2}Du+uDe^{x^2}$, $D^2y = e^{x^2}D^2u+2Du.De^{x^2}+uD^2e^{x^2}$ and so $y = ue^{x^2}$ is a solution provided $e^{x^2}D^2u+2Du.De^{x^2}+uD^2e^{x^2} = 2(1+2x^2)ue^{x^2} = uD^2e^{x^2}$ (since e^{x^2} is a solution), i.e. $D^2u+4xDu = 0$; write $v = Du$, then $Dv+4xv = 0$, i.e. $(1/v)Dv+4x = 0$, whence $\log v = \log a-2x^2$, $v = ae^{-2x^2}$, and so $u = a\int e^{-2x^2}dx+b$, whence $y = ae^{x^2}\int e^{-2x^2}dx+be^{x^2}$. The integral $\int e^{-2x^2}dx$ cannot be expressed in terms of any combination of circular or exponential functions; its value is given by integrating the series

$$1-(2x)^2/1!+(2x)^4/2!-(2x)^6/3!+\ldots$$

so that $$\int_0^x e^{-x^2}\,dx = x-(2x)^3/1!\,6+(2x)^5/2!\,10-(2x)^7/3!\,14+\ldots.$$

Alternatively the second solution may be obtained as follows; let $y = y_1$ and $y = y_2$ be two solutions, so that

$$d^2y_1/dx^2 = 2(1+2x^2)y_1, \qquad d^2y_2/dx^2 = 2(1+2x^2)y_2,$$

whence $y_2\,d^2y_1/dx^2 - y_1\,d^2y_2/dx^2 = 0$; write $W = y_2\,dy_1/dx - y_1\,dy_2/dx$, then $dW/dx = y_2\,d^2y_1/dx^2 - y_1\,d^2y_2/dx^2 = 0$ so that $W =$ constant $= a$; hence

$$D_x(y_1/y_2) = W/y_2^2 = a/y_2^2 = ae^{-2x^2},$$

taking for y_2 the solution e^{x^2}, and therefore

$$y_1 = ay_2 \int e^{-2x^2}\,dx + by_2 = ae^{x^2} \int e^{-2x^2}\,dx + be^{x^2}.$$

Since $\int e^{-2x^2}\,dx = (1/\sqrt{2}) \int e^{-y^2}\,dy$, $y = x\sqrt{2}$, the solution may be written in the form $y_1 = ae^{x^2} \int e^{-x^2}\,dx + be^{x^2}$.

11.6. Let $y = 1 + x^3/3! + x^6/6! + \dots$; the series is convergent for any value of x, therefore $Dy = x^2/2! + x^5/5! + \dots$ and $D^2y = x + x^4/4! + \dots$ so that $D^2y + Dy + y = \sum x^n/n! = e^x$, i.e. $(D^2+D+1)y = e^x$. When $x = 0$, $y = 1$ and $Dy = 0$. The general solution of $(D^2+D+1)y = e^x$ is

$$y = \tfrac{1}{3}e^x + e^{-\frac{1}{2}x}\{A \cos \tfrac{1}{2}\sqrt{3}\,x + B \sin \tfrac{1}{2}\sqrt{3}\,x\};$$

taking $x = 0$ in y and Dy, we find $1 = \frac{1}{3} + A$ and $0 = \frac{1}{3} - A/2 + B\sqrt{3}/2$, whence $A = \frac{2}{3}$, $B = 0$.

11.7. $kyx^{y-1} = \dfrac{d}{dx}x^y = x^y\left\{\dfrac{y}{x} + \dfrac{dy}{dx}\log x\right\}$ and so $(k-1)y = \dfrac{dy}{dx}x\log x$,

whence $$\int \frac{1}{y}\,dy = (k-1)\int \frac{1}{x\log x}\,dx + \log a,$$

i.e. $\log y = \log a + (k-1)\log\log x$ or $y = a(\log x)^{k-1}$.

11.8. $\qquad -D^2\phi/(D\phi)^2 = 1/\phi$, and so $D^2\phi/D\phi + D\phi/\phi = 0$,

whence $\log D\phi + \log \phi = \alpha$, i.e. $2\phi\,D\phi = a$ and so $\phi^2 = ax+b$.

11.9. The Wronskian of t_1, t_2 is

$$W(t_1,t_2) = \begin{vmatrix} e^{px} & xe^{px} \\ pe^{px} & pxe^{px}+e^{px} \end{vmatrix} = \begin{vmatrix} e^{px} & 0 \\ 0 & e^{px} \end{vmatrix}$$

$$W(u_1,u_2) = \begin{vmatrix} \sin ax & \cos ax \\ a\cos ax & -a\sin ax \end{vmatrix} = -a(\sin^2 ax + \cos^2 ax) = -a \neq 0.$$

$W(u_1,u_2,v_1,v_2)$

$$\begin{vmatrix} \sin ax & \cos ax & x\sin ax & x\cos ax \\ a\cos ax & -a\sin ax & ax\cos ax + \sin ax & -ax\sin ax + \cos ax \\ -a^2\sin ax & -a^2\cos ax & -a^2x\sin ax + 2a\cos ax & -a^2x\cos ax - 2a\sin ax \\ -a^3\cos ax & a^3\sin ax & -a^3x\cos ax - 3a^2\sin ax & a^3x\sin ax - 3a^2\cos ax \end{vmatrix}$$

$$= \begin{vmatrix} \sin ax & \cos ax & 0 & 0 \\ a\cos ax & -a\sin ax & \sin ax & \cos ax \\ -a^2\sin ax & -a^2\cos ax & 2a\cos ax & -2a\sin ax \\ -a^3\cos ax & a^3\sin ax & -3a^2\sin ax & -3a^2\cos ax \end{vmatrix}$$

$$= \begin{vmatrix} \sin ax & \cos ax & 0 & 0 \\ 0 & 0 & \sin ax & \cos ax \\ a^2 \sin ax & a^2 \cos ax & 2a \cos ax & -2a \sin ax \\ 2a^3 \cos ax & -2a^3 \sin ax & -3a^2 \sin ax & -3a^2 \cos ax \end{vmatrix}$$

$$= \begin{vmatrix} \sin ax & \cos ax & 0 & 0 \\ 0 & 0 & \sin ax & \cos ax \\ 0 & 0 & 2a \cos ax & -2a \sin ax \\ 2a^3 \cos ax & -2a^3 \sin ax & -3a^2 \sin ax & -3a^2 \cos ax \end{vmatrix}$$

$$= 4a^4 \begin{vmatrix} \sin ax & \cos ax \\ \cos ax & -\sin ax \end{vmatrix}^2 = 4a^4 \neq 0.$$

11.91. If $y_1, y_2, \ldots, y_n$ is a fundamental set of solutions of $L(D)y = 0$, then any other solution η is such that $\eta = A_1 y_1 + A_2 y_2 + \ldots + A_n y_n$, where each A_r is constant. The Wronskian of $y_1, y_2, \ldots, y_n$ is not zero for any x, in particular for $x = a$, so that the set of equations $D^r \eta = \sum_k A_k D^r y_k$, $0 \leqslant r \leqslant n-1$, is solvable at $x = a$. Thus $A_1, A_2, \ldots, A_n$ are uniquely determined by the value of $D^r \eta$ at $x = a$, $0 \leqslant r \leqslant n-1$, and therefore η is uniquely determined.

If $D^r \eta = 0$, at $x = a$, $0 \leqslant r \leqslant n-1$, then $\sum_k A_k D^r y_k = 0$, at $x = a$, $0 \leqslant r \leqslant n-1$. Since the determinant of the coefficients of the A_k's is not zero, the only solution of the set of equations $\sum_k A_k D^r y_k = 0$, $0 \leqslant r \leqslant n-1$, is $A_1 = A_2 = \ldots = A_n = 0$, and therefore $\eta = 0$ for all values of x.

11.92. Since the determinant $W(\eta_1, \eta_2, \ldots, \eta_n)$ is zero at $x = a$, the set of equations $\sum_k A_k D^r \eta_k = 0$, $0 \leqslant r \leqslant n-1$, is solvable for $A_1, A_2, \ldots, A_k$, not all zero, at $x = a$. Then $\eta = \sum A_k \eta_k$ is a solution of $L(D)y = 0$ such that, at $x = a$, $\eta = D\eta = D^2\eta = \ldots = D^{n-1}\eta = 0$, and therefore, by Example 11.91, $\eta = 0$ for all values of x. Differentiating the equation $\eta = 0$ repeatedly, we find $\sum_k A_k D^r \eta_k = 0$, $0 \leqslant r \leqslant n-1$. Since the A's are not all zero, it follows that the determinant of the coefficients of the A's is zero, for all values of x, i.e. $W(\eta_1, \eta_2, \ldots, \eta_n)$ is zero for all values of x.

11.93. Take $D-1$ times the second equation from the first, giving

$$(D^2+4)v = \sin x,$$

of which the general solution is $v = A \sin 2x + B \cos 2x + \frac{1}{3} \sin x$. Whence, from the second equation,

$$(D+1)u = 2A \cos 2x - 2B \sin 2x + \tfrac{1}{3} \cos x + x^2,$$

and the general solution of this is

$$u = Ce^{-x} + \tfrac{2}{5}(A+2B) \cos 2x + \tfrac{2}{5}(2A-B) \sin 2x + \tfrac{1}{6}(\sin x + \cos x) + x^2 - 2x + 2;$$

it is readily verified that the first equation is also satisfied.

11.94. The formula holds with $n = 0$ for all p. If it is true for all p with $n = 1, 2, \ldots, k$ then

$$(D^2+a^2)^{k+1}x^{k+p+1}\sin ax = (D^2+a^2)^k\{(D^2+a^2)x^{k+p+1}\sin ax\}$$

$$= (D^2+a^2)^k\{2a(k+p+1)x^{k+p}\cos ax+(k+p+1)(k+p)x^{k+p-1}\sin ax\}$$

$$= 2a\{(k+p+1)!\}\left\{\frac{x^p}{p!}(2D)^k+\frac{x^{p-1}}{(p-1)!}k.(2D)^{k-1}+\ldots\right\}\cos ax+$$

$$+(k+p+1)!\left\{\frac{x^{p-1}}{(p-1)!}(2D)^k+\ldots\right\}\sin ax$$

$$= (k+p+1)!\left[\frac{x^p}{p!}(2D)^{k+1}+(k+1)\frac{x^{p-1}}{(p-1)!}(2D)^k+\left\{\binom{k}{2}+\binom{k}{1}\right\}\frac{x^{p-2}}{(p-2)!}(2D)^{k-1}+\right.$$

$$\left.+\left\{\binom{k}{3}+\binom{k}{2}\right\}\frac{x^{p-3}}{(p-3)!}(2D)^{k-2}+\ldots\right]\sin ax$$

$$= (k+p+1)!\left\{\frac{x^p}{p!}(2D)^{k+1}+\binom{k+1}{1}\frac{x^{p-1}}{(p-1)!}(2D)^k+\binom{k+1}{2}\frac{x^{p-2}}{(p-2)!}(2D)^{k-1}+\right.$$

$$\left.+\binom{k+1}{3}\frac{x^{p-3}}{(p-3)!}(2D)^{k-2}+\ldots\right\}\sin ax.$$

Since we may replace $\sin ax$ by $\cos ax$, the formula holds also for $n = k+1$, and so by induction for all values of n.

XII

12. By Rolle's theorem $f'(x)$ has at least one root between any two of $f(x)$, so that $f'(x)$ has at least $n-1$ roots; hence $f''(x)$ has at least $n-2$ roots, $f'''(x)$ at least $n-3$, and so on.

12.01. The numbers a, b, c are the roots of the equation

$$f(t) = t^3-2t^2+t-abc = t(t-1)^2-abc = 0.$$

But $f'(t) = (3t-1)(t-1)$, and a root of $f'(t) = 0$ lies between two of the roots of $f(t) = 0$, by Rolle's theorem, and therefore $a < \frac{1}{3} < b < 1 < c$.

Since $f(c) = 0$, therefore $ab = (c-1)^2$ so that $a > 0$, and therefore $f(0) = -abc < 0$. Since $f'(\frac{1}{3}) = 0$, $f(t)-f(\frac{1}{3})$ has the factor $(t-\frac{1}{3})^2$, and since the sum of the roots of the equation $f(t)-f(\frac{1}{3}) = 0$ is 2, the third root is $t = \frac{4}{3}$. Thus $f(t) = (t-\frac{4}{3})(t-\frac{1}{3})^2+f(\frac{1}{3})$. One root only of $f(t) = 0$ lies between 0 and $\frac{1}{3}$ and $f(0) < 0$, therefore $f(\frac{1}{3}) > 0$. But $f(c) = 0$ and so $(c-\frac{4}{3})(c-\frac{1}{3})^2 = -f(\frac{1}{3}) < 0$. Thus $c < \frac{4}{3}$. Hence

$$0 < a < \tfrac{1}{3} < b < 1 < c < \tfrac{4}{3}.$$

12.1. By the mean-value theorem we can find c_x such that

$$f(x) = (x-\lambda)f'(c_x);$$

let $g(x) = f'(c_x)$, $x \neq \lambda$, $g(\lambda) = f'(\lambda)$, then, by Example 3.5, $g(x)$ is continuous in (a, b). It does not follow that $g(x)$ is differentiable; e.g. $(x-\lambda)^{\frac{4}{3}} = (x-\lambda)(x-\lambda)^{\frac{1}{3}}$ but $(x-\lambda)^{\frac{1}{3}}$ is not differentiable in an interval which contains λ.

12.11. Immediate consequence of Examples 3.5 and 3.6.

12.12. Let

$$H(t) = g(x)(t-a)(t-b)-g(t)(x-a)(x-b), \quad a < x < b,$$

then $H(t)$ vanishes at $t = a$, $t = x$, and $t = b$ and so, by Rolle's theorem, $H'(t)$ vanishes between a and x, and again between x and b, so that $H''(t)$ vanishes (at least) once at c, say, in $[a, b]$.

But $$H''(t) = 2g(x)-(x-a)(x-b)g''(t)$$

and so $$g(x) = \tfrac{1}{2}(x-a)(x-b)g''(c),$$

which proves that $g(x) \neq 0$ in $[a, b]$, since $g''(c) \neq 0$.

12.2. If

$$\{f(X)-f(a)\}/\{g(X)-g(a)\} > \{f(b)-f(a)\}/\{g(b)-g(a)\}$$

then $$f(X) > f(a)+\{f(b)-f(a)\}\{g(X)-g(a)\}/\{g(b)-g(a)\}$$

and so $$f(b)-f(X) < \{f(b)-f(a)\}\{g(b)-g(X)\}/\{g(b)-g(a)\},$$

whence $$\nu(X, b) < \nu(a, b), \quad \text{etc.}$$

12.21. $$\{f(x)-f(a)\}/\{g(x)-g(a)\} = \nu(a, x) = \nu(a, b)$$

and so $$f(x)-f(a) = \{g(x)-g(a)\}\nu(a, b)$$

whence $$f'(x) = g'(x)\nu(a, b) \quad \text{for all } x \text{ in } (a, b).$$

12.22. Since $g(x)$ is differentiable we can determine q depending on p so that $\{g(y)-g(x)\}/(y-x) = g'(x)+0(p)$ provided $y-x = 0(q)$. Choose p so that $0(p) < \lambda$ then $\{g(y)-g(x)\}/(y-x) \geqslant \lambda$ and so, if $y > x$,

$$g(y)-g(x) \geqslant \lambda(y-x), \quad \text{provided } y-x = 0(q).$$

If α and β are any two points in (a, b), divide (α, β) into k equal parts (α_r, α_{r+1}), $r = 0, 1, 2,\ldots, k-1$, $\alpha_0 = \alpha$, $\alpha_k = \beta$, so that $\alpha_{r+1}-\alpha_r = 0(q)$, then

$$g(\beta)-g(\alpha) = \sum_0^{k-1}\{g(\alpha_{r+1})-g(\alpha_r)\} \geqslant \lambda\sum_0^{k-1}(\alpha_{r+1}-\alpha_r) = \lambda(\beta-\alpha), \text{ i.e.}$$

$$g(\beta)-g(\alpha) \geqslant \lambda(\beta-\alpha).$$

Divide the interval $g(a)$, $g(b)$ into two equal parts by the point A. Since $g(a) < A < g(b)$ and $g(x)$ is continuous we can find $\bar{a}$ so that $g(\bar{a}) = A$. Now

$$\{f(\bar{a})-f(a)\}/\{g(\bar{a})-g(a)\}+\{f(b)-f(\bar{a})\}/\{g(b)-g(\bar{a})\}$$
$$= \{f(b)-f(a)\}/\tfrac{1}{2}\{g(b)-g(a)\},$$

i.e. $$\nu(\bar{a}, a)+\nu(b, \bar{a}) = 2\nu(a, b).$$

Hence $\nu(a, b) \leqslant \max\{\nu(\bar{a}, a), \nu(b, \bar{a})\}$ and $\nu(a, b) \geqslant \min\{\nu(\bar{a}, a), \nu(b, \bar{a})\}$. Let (a_1, b_1) denote that of the two intervals $(a, \bar{a})$, $(b, \bar{a})$ for which

$$\nu(a, b) \leqslant \nu(a_1, b_1).$$

Next divide $g(a_1)$, $g(b_1)$ into two equal parts, and choose that part $g(a_2)$, $g(b_2)$, say, for which $\nu(a_1, b_1) \leqslant \nu(a_2, b_2)$, and so on. Thus we have determined intervals (a_n, b_n) each of which is contained in its predecessor and such that $\nu(a, b) \leqslant \nu(a_n, b_n)$. But $b_n-a_n \leqslant \{g(b_n)-g(a_n)\}/\lambda = \{g(b)-g(a)\}/2^n\lambda \to 0$, and a_n and b_n tend to a common limit c_2, say; hence

$$\nu(a_n, b_n) = [\{f(b_n)-f(a_n)\}/(b_n-a_n)]/[\{g(b_n)-g(a_n)\}/(b_n-a_n)]$$

tends to $f'(c_2)/g'(c_2)$ and so $\nu(a,b) \leqslant f'(c_2)/g'(c_2)$. Similarly we can determine c_1 so that $\nu(a,b) \geqslant f'(c_1)/g'(c_1)$.

12.23. Since $\nu(a,x)$ is not constant we can find X, $a < X < b$, such that $\nu(a,X)$ is different from $\nu(a,b)$; suppose $\nu(a,X) > \nu(a,b)$, then

$$\nu(X,b) < \nu(a,b).$$

By 12.22 we can determine a c_2 in (a,X) such that

$$f'(c_2)/g'(c_2) \geqslant \nu(a,X) > \nu(a,b)$$

and a c_1 in (X,b) such that $f'(c_1)/g'(c_1) \leqslant \nu(X,b) < \nu(a,b)$. Thus $\nu(a,b)$ lies between the values of $f'(x)/g'(x)$ at c_1 and c_2. Since $g'(x) \geqslant 2\lambda > 0$, $f'(x)/g'(x)$ is continuous and so there is a point c, between c_1 and c_2, such that $\nu(a,b) = f'(c)/g'(c)$.

12.3. Let
$$\phi(h) = \int_{-h}^{h} f(x)\,dx - 2hf(0),$$

then
$$\phi(0) = 0, \qquad \phi'(h) = f(h)+f(-h)-2f(0),$$
$$\phi'(0) = 0, \qquad \phi''(h) = f'(h)-f'(-h) = 2hf''(c_h),$$

by the mean-value theorem. Hence by Theorem 12.52 there is a point u in $(0,h)$ such that $\dfrac{\phi(h)}{h^3} = \dfrac{\phi''(u)}{3.2.u} = \frac{1}{3}f''(c_u)$ and so $\phi(h) = \frac{1}{3}h^3f''(\alpha)$, where $\alpha = c_u$, which is a point in $[-u,u]$ and so a point in $[-h,h]$.

12.31. If $\phi(h) = \int_{-h}^{h} f(x)\,dx - h\{f(h)+f(-h)\}$ then

$$\phi'(h) = -h\{f'(h)-f'(-h)\} = -2h^2f''(c_h);$$

since $\phi(0) = 0$,
$$\frac{\phi(h)}{h^3} = \frac{\phi'(u)}{3u^2} = -\tfrac{2}{3}f''(c_u),$$

whence $\phi(h) = -\frac{2}{3}h^3f''(\beta)$, where $\beta = c_u$.

12.32. If
$$\phi(h) = \int_{-h}^{h} f(x)\,dx - (h/2)\{f(h)+2f(0)+f(-h)\}$$

then
$$\phi(0) = \phi'(0) = 0 \quad \text{and} \quad \phi''(h) = -(h/2)\{f''(h)+f''(-h)\};$$

but $\{f''(h)+f''(-h)\}/2$ lies between $f''(h)$ and $f''(-h)$ and so, since $f''(x)$ is continuous, equals $f''(c)$ for a certain c in $(-h,h)$. Thus $\phi''(h) = -hf''(c)$ and so $\phi(h) = -(h^3/6)f''(\gamma)$.

12.33. Let

$$12\phi(h) = 12\int_{0}^{h} f(x)\,dx - h\{-f(-h)+8f(0)+5f(h)\}$$

then $\phi(0) = \phi'(0) = \phi''(0) = 0$ and
$$12\phi'''(h) = -h\{f'''(-h)+5f'''(h)\} - 3\{f''(h)-f''(-h)\}$$
$$= -h\{f'''(-h)+5f'''(h)+6f'''(c)\}$$

by the mean-value theorem.

But

$$\min\{f'''(-h), f'''(h), f'''(c)\} \leqslant \{f'''(-h)+5f'''(h)+6f'''(c)\}/12$$
$$\leqslant \max\{f'''(-h), f'''(h), f'''(c)\}$$

and so, since $f'''(x)$ is continuous, there is a point c_1 between $-h$ and h such that $\{f'''(-h)+5f'''(h)+6f'''(c)\}/12 = f'''(c_1)$; thus $\phi'''(h) = -hf'''(c_1)$ and so $\phi(h) = -(h^4/4!)f'''(\delta)$.

12.331. Write $\phi(h) = f(a+h)-f(a-h)-2hf'(a)$, then

$$\phi(0) = \phi'(0) = \phi''(0), \quad \text{and} \quad \phi''(h) = f''(a+h)-f''(a-h) = 2hf^3(\xi)$$

and so, by Theorem 12.52,

$$\frac{\phi(h)}{h^3} = \frac{\phi(h)-\phi(0)}{h^3-0^3} = \frac{\phi''(c)}{6c} = \tfrac{1}{3}f^3(\alpha), \quad \alpha = \xi(c).$$

12.332. Write $G(x) = g(x+h)-g(x-2h)$, then

$$G'(x) = g'(x+h)-g'(x-2h) = 3hg''(x+\theta h), \quad -2 < \theta < 1,$$

and so

$$G(a+h)-G(a) = hG'(a+\phi h) = 3h^2g''(c), \quad a-2h < c < a+2h.$$

12.333. Write

$$\phi(h) = f(a+2h)-8f(a+h)+8f(a-h)-f(a-2h)+12hf'(a)$$

then
$$\phi(0) = \phi'(0) = \phi''(0) = 0$$

and
$$\phi'''(h) = 8\{f'''(a+2h)-f'''(a+h)-f'''(a-h)+f'''(a-2h)\}$$
$$= 24h^2f^5(c), \quad \text{by Example 12.332.}$$

Therefore $$\frac{\phi(h)}{h^5} = \frac{\phi'''(\alpha)}{60\alpha^2} = \tfrac{2}{5}f^5(\beta), \quad \text{by Theorem 12.52.}$$

12.334. Write

$$\phi(h) = f(a+h)-2f(a)+f(a-h)-h^2f''(a)$$

so that
$$\phi(0) = \phi'(0) = \phi''(0) = 0,$$

and
$$\phi'''(h) = f'''(a+h)-f'''(a-h) = 2hf^4(\xi)$$

and therefore
$$\frac{\phi(h)}{h^4} = \frac{\phi'''(\alpha)}{24\alpha} = \frac{f^4(\beta)}{12}.$$

12.335. Write

$$\phi(h) = f(a+2h)-16f(a+h)+30f(a)-16f(a-h)+f(a-2h)+12h^2f''(a)$$

so that
$$\phi(0) = \phi'(0) = \phi''(0) = \phi'''(0) = 0$$

and
$$\phi^4(h) = 16\{f^4(a+2h)-f^4(a+h)-f^4(a-h)+f^4(a-2h)\}$$
$$= 16.3h^2f^6(\xi), \quad \text{by Example 12.332,}$$

and therefore
$$\frac{\phi(h)}{h^6} = \frac{\phi^4(\alpha)}{6.5.4.3.\alpha^2} = \frac{2}{15}f^6(\beta).$$

12.4. By the generalized integral mean-value theorem

$$\int_n^N f(x)g(x)\,dx = g(n)\int_n^\nu f(x)\,dx + g(N)\int_\nu^N f(x)\,dx, \quad n < \nu < N.$$

Since $f(x) > 0$ and $\int_a^\infty f(x)\,dx$ exists, therefore for $n \geqslant n_k$,

$$\int_n^\nu f(x)\,dx < \int_n^N f(x)\,dx < 1/k, \qquad \int_\nu^N f(x)\,dx < \int_n^N f(x)\,dx < 1/k$$

and since $g(x)$ is bounded, $|g(n)| \leqslant M$, $|g(N)| \leqslant M$, and so

$$\left|\int_n^N f(x)g(x)\,dx\right| \leqslant 2M/k,$$

which proves $\int^\infty f(x)g(x)\,dx$ converges.

12.41. As above,

$$\int_n^N f(x)g(x)\,dx = g(n)\int_n^\nu f(x)\,dx + g(N)\int_\nu^N f(x)\,dx < 2M\{|g(n)|+|g(N)|\} \to 0$$

12.5. N and n are integers such that $Na > b/n$, then

$$\int_{1/n}^N \{f(ax)-f(bx)\}\,dx/x = \int_{1/n}^N f(ax)\,dx/x - \int_{1/n}^N f(bx)\,dx/x$$

$$= \int_{a/n}^{Na} f(t)\,dt/t - \int_{b/n}^{Nb} f(u)\,du/u, \quad t = ax,\ u = bx,$$

$$= \int_{a/n}^{b/n} f(x)\,dx/x - \int_{Na}^{Nb} f(x)\,dx/x$$

$$= f(\alpha/n)\int_{a/n}^{b/n} dx/x - f(N\beta)\int_{Na}^{Nb} dx/x, \quad a < \alpha < b,\ a < \beta < b,$$

$$= \{f(\alpha/n)-f(N\beta)\}\log b/a$$

$$\to (\lambda-\mu)\log b/a, \quad \text{since } \alpha/n \to 0 \text{ and } 1/N\beta \to 0.$$

12.51.
$$\int_{1/n}^N \{f(ax)-f(bx)\}\,dx/x = f(\alpha/n)\log b/a - \int_{Na}^{Nb} f(x)\,dx/x$$

$$= f(\alpha/n)\log b/a - \int_a^b f(Nt)\,dt/t, \quad x = Nt,$$

$$\to \lambda \log b/a.$$

12.52.
$$\int_{1/n}^N \{f(ax)-f(bx)\}\,dx/x = \int_{a/n}^{b/n} f(x)\,dx/x - f(N\beta)\log b/a$$

$$= \int_a^b f(u/n)\,du/u - f(N\beta)\log b/a, \quad u = nx.$$

$$\to \mu \log a/b.$$

12.6. The sequence $\int\limits_{1/n}^{X} D_x(1/\sin x - 1/x)\cos ax\, dx$ converges, for if $N > n$,

$$\left|\int\limits_{1/N}^{X} - \int\limits_{1/n}^{X} \{D_x(1/\sin x - 1/x)\cos ax\}\, dx\right| = \left|\int\limits_{1/N}^{1/n} D_x(1/\sin x - 1/x)\cos ax\, dx\right|$$

$$= \left|\cos a/\nu \int\limits_{1/N}^{1/n} D_x(1/\sin x - 1/x)\, dx\right|, \quad n < \nu < N,$$

since $D_x(1/\sin x - 1/x) > 0$, by Example 5.31,

$$= \left|\cos a/\nu [1/\sin x - 1/x]_{1/N}^{1/n}\right|$$

$$< 1/\sin(1/n) - n < n/(6n^2 - 1), \quad \text{by Example 5.32.}$$

Furthermore

$$\left|\int\limits_{1/n}^{X} D_x(1/\sin x - 1/x)\cos ax\, dx\right|$$

$$= |\cos aX_n[\{1/\sin X - 1/X\} - \{1/\sin(1/n) - n\}]|, \quad 1/n < X_n < X,$$

$$< 1/\sin X - 1/X.$$

In particular $\left|\int\limits_{0}^{\frac{1}{2}\pi} D_x(1/\sin x - 1/x)\cos ax\, dx\right| \leqslant (\pi - 2)/\pi.$

12.61. $$\int\limits_{0}^{\frac{1}{2}\pi} \frac{\sin(2n+1)x}{\sin x}\, dx = \int\limits_{0}^{\frac{1}{2}\pi} \{1 + 2\cos 2x + 2\cos 4x + \dots + 2\cos 2nx\}\, dx,$$

by Example 5.12,

$$= \tfrac{1}{2}\pi.$$

Furthermore

$$\left|\int\limits_{0}^{\frac{1}{2}\pi} \left\{\frac{\sin(2n+1)x}{\sin x} - \frac{\sin(2n+1)x}{x}\right\} dx\right|$$

$$= \left|\int\limits_{0}^{\frac{1}{2}\pi} \sin(2n+1)x(1/\sin x - 1/x)\, dx\right|$$

$$= \left|\frac{1}{2n+1} \int\limits_{0}^{\frac{1}{2}\pi} D_x(1/\sin x - 1/x)\cos(2n+1)x\, dx\right|$$

$$< (1 - 2/\pi)/(2n+1) \to 0, \quad \text{by Example 12.6.}$$

Thus $\int\limits_{0}^{\frac{1}{2}\pi} \frac{\sin(2n+1)x}{x}\, dx \to \frac{1}{2}\pi$; but

$$\int\limits_{0}^{\frac{1}{2}\pi} \frac{\sin(2n+1)x}{x}\, dx = \int\limits_{0}^{(n+\frac{1}{2})\pi} \frac{\sin t}{t}\, dt, \quad t = (2n+1)x,$$

and so $$\int\limits_{0}^{\infty} \frac{\sin t}{t}\, dt = \tfrac{1}{2}\pi.$$

12.62. $$\int_{1/n}^{N} \frac{\sin x}{x}\,dx = \left[\frac{1-\cos x}{x}\right]_{1/n}^{N} + \int_{1/n}^{N} \frac{1-\cos x}{x^2}\,dx,$$

integrating by parts,

$$\to \int_{0}^{\infty} \frac{1-\cos x}{x^2}\,dx$$

since $$\left|\frac{1-\cos N}{N}\right| \leqslant \frac{2}{N} \to 0$$

and $$0 < \frac{1-\cos 1/n}{1/n} < n\{1-(1-1/2n^2)\} = 1/2n \to 0$$

Furthermore

$$\int_{1/n}^{N} \frac{1-\cos x}{x^2}\,dx = \int_{1/n}^{N} \frac{2\sin^2\frac{1}{2}x}{x^2}\,dx = \int_{1/2n}^{N/2} \frac{\sin^2 t}{t^2}\,dt \to \int_{0}^{\infty} \left(\frac{\sin t}{t}\right)^2 dt.$$

12.63. $$\int_{1/n}^{N} \frac{\sin ax}{x}\,dx = \int_{a/n}^{Na} \frac{\sin t}{t}\,dt \to \int_{0}^{\infty} \frac{\sin t}{t}\,dt \quad \text{if} \quad a > 0$$
$$= \tfrac{1}{2}\pi.$$

If $a = -b$, $b > 0$, then $$\int_{1/n}^{N} \frac{\sin ax}{x}\,dx = -\int_{1/n}^{N} \frac{\sin bx}{x}\,dx \to -\tfrac{1}{2}\pi.$$

12.64. By the mean-value theorem

$$\log(x+1)-\log x = \frac{1}{c}, \quad 0 < x < c < x+1;$$

hence

$$D[x\{\log(x+1)-\log x\}] = \log(x+1)-\log x-\frac{1}{x+1} = \frac{1}{c}-\frac{1}{x+1} > 0,$$

and

$$D[(x+1)\{\log(x+1)-\log x\}] = \log(x+1)-\log x-\frac{1}{x} = \frac{1}{c}-\frac{1}{x} < 0,$$

so that $\left(1+\dfrac{1}{x}\right)^x$ is increasing and $\left(1+\dfrac{1}{x}\right)^{x+1}$ is decreasing for $x > 0$.

If $y = x+1$, then $\left(1+\dfrac{1}{x}\right)^{x+1} = \left(1-\dfrac{1}{y}\right)^{-y}$, so that $\left(1-\dfrac{1}{y}\right)^{-y}$ decreases for $y > 1$. Now $\left(1+\dfrac{1}{x}\right)^x < \left(1+\dfrac{1}{x}\right)^{x+1} < (1+1)^2$, $x > 1$, since $\left(1+\dfrac{1}{x}\right)^{x+1}$ decreases, and therefore

$$\left(1+\frac{1}{x}\right)^{x+1}-\left(1+\frac{1}{x}\right)^{x} = \left(1+\frac{1}{x}\right)^{x}\cdot\frac{1}{x} < \frac{4}{x} \to 0 \quad \text{as} \quad x \to \infty.$$

Hence if $a_n = \left(1+\dfrac{1}{n}\right)^n$, $b_n = \left(1+\dfrac{1}{n}\right)^{n+1}$ then a_n is steadily increasing, b_n is steadily decreasing, and $a_n < b_n$ and $b_n-a_n \to 0$. Thus (a_n, b_n) forms

a nest of intervals which contain a unique point which is the limit of both a_n and b_n. But if $x > n$, then

$$a_n = \left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{x}\right)^x < \left(1+\frac{1}{x}\right)^{x+1} < \left(1+\frac{1}{n}\right)^{n+1} = b_n$$

and so both $\left(1+\frac{1}{x}\right)^x$ and $\left(1+\frac{1}{x}\right)^{x+1}$ tend to the common limit of a_n and b_n.

12.7. Let $10\phi(h)$

$$= 3h[f(-3h)+5f(-2h)+f(-h)+6f(0)+f(h)+5f(2h)+f(3h)]- \\ -10\int_{-3h}^{3h} f(x)\,dx$$

then

$$\phi(0) = \phi'(0) = \phi''(0) = \phi'''(0) = \phi^{\mathrm{iv}}(0) = \phi^{\mathrm{v}}(0) = \phi^{\mathrm{vi}}(0) = 0$$

and

$$\begin{aligned}10\phi^{\mathrm{vi}}(h) &= 18[243f^{\mathrm{v}}(3h)+160f^{\mathrm{v}}(2h)+f^{\mathrm{v}}(h)-f^{\mathrm{v}}(-h)-160f^{\mathrm{v}}(-2h)- \\ &\quad -243f^{\mathrm{v}}(-3h)]+3h[729f^{\mathrm{vi}}(3h)+320f^{\mathrm{vi}}(2h)+f^{\mathrm{vi}}(h)+f^{\mathrm{vi}}(-h)+ \\ &\quad +320f^{\mathrm{vi}}(-2h)+729f^{\mathrm{vi}}(-3h)]-7290[f^{\mathrm{v}}(3h)-f^{\mathrm{v}}(-3h)] \\ &= 18[2hf^{\mathrm{vi}}(\alpha)+160.4hf^{\mathrm{vi}}(\beta)-162.6hf^{\mathrm{vi}}(\gamma)]+3h[729\{f^{\mathrm{vi}}(3h)+ \\ &\quad +f^{\mathrm{vi}}(-3h)\}+320\{f^{\mathrm{vi}}(2h)+f^{\mathrm{vi}}(-2h)\}+\{f^{\mathrm{vi}}(h)+f^{\mathrm{vi}}(-h)\}], \\ &\quad -h < \alpha < h,\ -2h < \beta < 2h,\ -3h < \gamma < 3h,\end{aligned}$$

and so

$$\begin{aligned}10\phi^{\mathrm{vi}}(h) &< \{18(2+640+972)+6(730+320)\}Mh \\ &= 35352Mh,\end{aligned}$$

whence, integrating six times from 0 to h, we have

$$\phi(h) < \{35352/10(7!)\}Mh^7 = (491/700)Mh^7 < \tfrac{5}{7}Mh^7.$$

12.8.
$$\int_0^3 f(x)g(x)\,dx = \int_0^3 (x-1)^{\frac{4}{3}}\,dx = \tfrac{3}{7}(4\sqrt[3]{2}+1);$$

$$f(X)\int_0^3 g(x)\,dx = (X-1)^{\frac{1}{3}}\int_0^3 (x-1)\,dx = \tfrac{3}{2}(X-1)^{\frac{1}{3}},$$

and, in $(0,3)$, $\quad \tfrac{3}{2}(X-1)^{\frac{1}{3}} \leqslant \tfrac{3}{2}\sqrt[3]{2} < 3.\tfrac{4}{7}\sqrt[3]{2} < \tfrac{3}{7}(4\sqrt[3]{2}+1).$

12.81.
$$\int_0^4 f(x)g(x)\,dx = \int_0^4 (1-x)\sqrt{x}\,dx = -112/15,$$

and
$$f(0)\int_0^X g(x)\,dx = \tfrac{2}{3}X^{\frac{3}{2}} > -112/15, \quad X \geqslant 0.$$

12.82. By the mean-value theorem

$$|f'(x)+1| = |f'(x)-f'(0)| = |xf''(c)| < 2|a|/4|a| = \tfrac{1}{2}, \quad -2a < x < 2a.$$

Hence, since $f(a)-a = f(a)-f(0) = af'(\alpha)$, therefore

$$|f(a)| = |a|\,|1+f'(\alpha)| < \tfrac{1}{2}|a|.$$

Furthermore $|a+b| = |a+f(a)| < \frac{3}{2}|a|$ so that $a+b$ lies in $(-2a, 2a)$ and

so
$$f(a+b) = f(a)+bf'(\beta), \quad \beta \text{ in } (-2a, 2a),$$
$$= f(a)(1+f'(\beta))$$
and so
$$|f(a+b)| < \tfrac{1}{2}|f(a)| < \tfrac{1}{4}|a|.$$

XIII

13. The first five derivatives of $\tan x$ are $1+t^2$, $2t(1+t^2)$, $2(1+4t^2+3t^4)$, $8t(2+5t^2+3t^4)$, and $8(2+17t^2+30t^4+15t^6)$, where t stands for $\tan x$, and the values these take for $x = 0$ are 1, 0, 2, 0, 16, whence result follows from Theorem 13.521.

Since $(1-x^2)^{-\frac{1}{2}} = 1+x^2/2+3x^4/8+\dots$ for $|x| < 1$, therefore
$$\sin^{-1}x = x+x^3/6+3x^5/40+\dots$$
and so the values of the successive derivatives of $\sin^{-1}x$ at $x = 0$ are 1, 0, 1, 0, 9,..., whence the result follows from Theorem 13.521.

13.01.
$$\lim_{x\to 0}(1/\sin x-1/x) = \lim_{x\to 0}\{(x-\sin x)/x\sin x\}$$
$$= \lim_{x\to 0}\sin x/(2\cos x-x\sin x) = 0.$$

$$\lim_{x\to 1}(1+\cos\pi x)/\tan^2\pi x = \lim_{x\to 1}(-\pi\sin\pi x/2\pi\tan\pi x\sec^2\pi x)$$
$$= \lim_{x\to 1}(-1/2\sec^3\pi x) = \tfrac{1}{2}.$$

$$\{\log(1+x)\}^3/(\tan x-\sin x) = \{x-x^2(\tfrac{1}{2}+\alpha_x)\}^3/[x+x^3(\tfrac{1}{3}+\beta_x)-\{x-x^3(\tfrac{1}{6}+\gamma_x)\}]$$
$$= \{1-x(\tfrac{1}{2}+\alpha_x)\}^3/(\tfrac{1}{2}+\beta_x+\gamma_x) \to 2 \quad \text{as} \quad x\to 0.$$

$$\lim_{x\to\pi/2}\cos 3x/(e^{2x}-e^{\pi}) = \lim_{x\to\pi/2}(-3\sin 3x/2e^{2x}) = \tfrac{3}{2}e^{-\pi}.$$

$$\{\sin x\sin^{-1}x-x^2\} = [\{x-x^3/6+x^5(\tfrac{1}{120}+\alpha_x)\}\{x+x^3/6+x^5(\tfrac{3}{40}+\beta_x)\}-x^2]$$
$$= x^6\{(\tfrac{1}{18}+\alpha_x+\beta_x)-(x^2/6)(\tfrac{1}{15}+\beta_x-\alpha_x)+x^4(\tfrac{1}{120}+\alpha_x)(\tfrac{3}{40}+\beta_x)\}$$
and
$$(\tan x\tan^{-1}x-x^2) = \{x+x^3/3+x^5(\tfrac{2}{15}+\gamma_x)\}\{x-x^3/3+x^5(\tfrac{1}{5}+\delta_x)\}-x^2$$
$$= x^6\{(\tfrac{2}{9}+\gamma_x+\delta_x)+(x^2/3)(\tfrac{1}{15}+\delta_x-\gamma_x)+x^4(\tfrac{2}{15}+\gamma_x)(\tfrac{1}{5}+\delta_x)\}.$$
Hence
$$\lim_{x\to 0}(\sin x\sin^{-1}x-x^2)/(\tan x\tan^{-1}x-x^2) = \tfrac{1}{18}/\tfrac{2}{9} = \tfrac{1}{4}.$$

$$\lim_{x\to\pi}\{\cot^2x-(x-\pi)^{-2}\} = \lim_{y\to 0}(\cot^2y-y^{-2}), \quad y = x-\pi,$$
$$= \lim_{y\to 0}\{1/\sin^2y-1/y^2\}-1$$
$$= -1+\lim_{y\to 0}\{(\tfrac{1}{3}+2\alpha_y)-y^2(\tfrac{1}{6}+\alpha_y)^2\}/\{1-y(\tfrac{1}{6}+\alpha_y)\}^2$$
$$= -1+\tfrac{1}{3} = -\tfrac{2}{3}.$$

$$1-3\sin x/x(2+\cos x) = (2x+x\cos x-3\sin x)/x(2+\cos x)$$
$$= x^5(\tfrac{1}{60}+\alpha_x-3\beta_x)/x[3-x^2(\tfrac{1}{2}+\gamma_x)].$$
Hence
$$1/x^4-3\sin x/x^5(2+\cos x) \to \tfrac{1}{60}/3 = \tfrac{1}{180}.$$

13.02. Since $e^y = 1+y(1+\alpha_y)$, where $\alpha_y \to 0$ as $y \to 0$, therefore

$$(e^y-1)/y \to 1 \text{ as } y \to 0,$$

and so $n(e^{z/n}-1) \to z$, where $y = z/n$, whence $n(\sqrt[n]{x}-1) \to \log x$, $x = e^z$.

13.03. By Taylor's theorem

$$e^x = 1+x+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\frac{x^{n+1}}{(n+1)!}e^{\theta x}, \quad 0 < \theta < 1,$$

and since $e^{\theta x}$ is positive, therefore $x^{n+1}e^{\theta x}$ is positive or negative according as x^{n+1} is positive or negative.

13.1. We have $F(a)+A = 0$, $F(b)+A+B = 0$ and so $A = -F(a)$, $B = F(a)-F(b)$. Each choice of $G(x)$ gives a different form to the remainder. For instance, we may take (i) $G(x) = (x-a)/(b-a)$ or

$$\text{(ii)} \quad G(x) = \sin(x-a)/\sin(b-a), \quad \text{etc.}$$

Since $F(x)+A+BG(x)$ vanishes for $x = a$ and $x = b$, by Rolle's theorem there is a point c in (a,b) where $F'(c)+BG'(c) = 0$.

In (i) $\quad F'(x) = -(b-x)^n f^{n+1}(x)/n! \quad$ and $\quad G'(x) = 1/(b-a)$

and so $\quad (b-c)^n f^{n+1}(c)/n! = F(a)/(b-a) \quad$ since $F(b) = 0$.

Thus $\quad F(a) = (b-a)(b-c)^n f^{n+1}(c)/n!$

i.e.

$$f(b) = f(a)+\frac{(b-a)}{1!}f'(a)+\frac{(b-a)^2}{2!}f''(a)+\ldots+$$
$$+\frac{(b-a)^n}{n!}f^n(a)+(b-a)\frac{(b-c)^n}{n!}f^{n+1}(c), \quad a < c < b.$$

In (ii) $G(x) = \sin(x-a)/\sin(b-a)$, then $G'(x) = \cos(x-a)/\sin(b-a)$ and so

$$f(b) = f(a)+\frac{(b-a)}{1!}f'(a)+\frac{(b-a)^2}{2!}f''(a)+\ldots+$$
$$+\frac{(b-a)^n}{n!}f^n(a)+\frac{\sin(b-a)}{\cos(c-a)}\frac{(b-c)^n}{n!}f^{n+1}(c).$$

13.11. Here

$$F'(x) = x^n f^{n+1}(b-x)/n!, \qquad A = -F(h), \qquad B = F(h)-F(0) = F(h)$$

and so $$\frac{c^n}{n!}f^{n+1}(b-c) = -F(h)G'(c);$$

but $$F(h) = f(b)-f(a)-hf'(a)-\ldots-\frac{h^n}{n!}f^n(a)$$

and therefore

$$f(b) = f(a)+hf'(a)+\frac{h^2}{2!}f''(a)+\ldots+\frac{h^n}{n!}f^n(a)-\frac{c^n f^{n+1}(b-c)}{G'(c)n!}.$$

13.12. Here

$$f(b) = f(a)+hf'(a)+\ldots+\frac{h^n}{n!}f^n(a)+\frac{(h-c)^n f^{n+1}(a+c)}{G'(c)n!}.$$

13.13. By Taylor's theorem we can find θ_1, θ_2 such that

$$f(a+h) = f(a)+hf'(a)+\frac{h^2}{2!}f''(a)+\ldots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(a)+\frac{h^n}{n!}f^n(a+\theta_1 h)$$

and

$$f(a+k) = f(a)+kf'(a)+\frac{k^2}{2!}f''(a)+\ldots+\frac{k^{n-1}}{(n-1)!}f^{n-1}(a)+\frac{k^n}{n!}f^n(a+\theta_2 k),$$

whence

$$f(a+h)+f(a+k) = 2f(a)+(h+k)f'(a)+\frac{h^2+k^2}{2!}f''(a)+\ldots+$$
$$+\frac{h^{n-1}+k^{n-1}}{(n-1)!}f^{n-1}(a)+\frac{h^nf^n(a+\theta_1 h)+k^nf^n(a+\theta_2 k)}{n!}.$$

Since h, k have the same sign, $\dfrac{h^nf^n(a+\theta_1 h)+k^nf^n(a+\theta_2 k)}{h^n+k^n}$ lies between $f^n(a+\theta_1 h)$ and $f^n(a+\theta_2 h)$, and therefore, as $f^n(x)$ is continuous, we can find λ between $a+\theta_1 h$, $a+\theta_2 k$ such that

$$[h^nf^n(a+\theta_1 h)+k^nf^n(a+\theta_2 k)]/(h^n+k^n) = f^n(\lambda),$$

whence the result follows.

13.14. If $\phi(t) = f(a+ht)-f(a+kt)$, then $\phi^r(t) = h^rf^r(a+ht)-k^rf^r(a+kt)$, and so $\phi^r(0) = (h^r-k^r)f^r(a)$; but

$$\phi(1) = \phi(0)+\phi'(0)+\frac{1}{2!}\phi''(0)+\ldots+\frac{1}{(n-1)!}\phi^{n-1}(0)+\frac{1}{n!}\phi^n(\theta),$$

therefore

$$f(a+h)-f(a+k) = (h-k)f'(a)+\frac{h^2-k^2}{2!}f''(a)+\ldots+\frac{h^{n-1}-k^{n-1}}{(n-1)!}f^{n-1}(a)+$$
$$+\frac{h^nf^n(a+\theta h)-k^nf^n(a+\theta k)}{n!}$$

13.2. We have

$$f(a+h) = f(a)+hf'(a)+\frac{h^2}{2!}f''(a)+\ldots+\frac{h^{n-1}}{(n-1)!}f^{n-1}(a)+$$
$$+\frac{h^n}{n!}f^n(a+\theta_n h),\quad 0<\theta_n<1,$$

and

$$f(a+h) = f(a)+\ldots+\frac{h^n}{n!}f^n(a)+\frac{h^{n+1}}{(n+1)!}f^{n+1}(a+\theta_{n+1}h),\quad 0<\theta_{n+1}<1,$$

and so

$$\frac{h}{n+1}f^{n+1}(a+\theta_{n+1}h) = f^n(a+\theta_n h)-f^n(a) = \theta_n hf^{n+1}(a+\theta\theta_n h),\quad 0<\theta<1,$$

by the mean-value theorem, and so

$$\theta_n = \{1/(n+1)\}f^{n+1}(a+\theta_{n+1}h)/f^{n+1}(a+\theta\theta_n h) \to 1/(n+1)\quad \text{as}\quad h\to 0,$$

since $f^{n+1}(x)$ is continuous and $f^{n+1}(a) \neq 0$.

13.21. Since $f(x)\to f(a)$ as $x\to a$, we can find α_k so that $|f(x)-f(a)| < 1/2k$ provided $a \leqslant x \leqslant \alpha_k$ and therefore

$$|f(X)-f(x)| = |f(X)-f(a)-\{f(x)-f(a)\}| < 1/k$$

for any x, X in the interval (a, α_k). But $f(x)$ is continuous in (α_k, b) so that (α_k, b) may be divided into a finite number of parts such that, for any x, X in the same part, $|f(X)-f(x)| < 1/k$. Thus we can divide (a, b) into a finite number of parts such that $|f(X)-f(x)| < 1/k$ for any x, X in the same part, which proves that $f(x)$ is continuous in (a, b).

13.3. If $f(x) = x+e^x-9\cdot 3$, $f(2) = \cdot 08906$, $f'(2) = 8\cdot 38906$, and so if $k = \frac{1}{8}$, $kf'(2) = 1\cdot 048...$; in the interval $2-2kf(2)$, $2+2kf(2)$, i.e. in $(1\cdot 98, 2\cdot 03)$, $f''(x) = e^x < 8 = M$ (say), and therefore

$$\cdot 08906 = f(2) < 1/12k^2M = \cdot\dot{6}.$$

Hence we may apply Theorem 13.93, and we have in turn

$c_1 = 2-f(2)/8 = 2-\cdot 01113 = 1\cdot 98887$, $c_2 = 1\cdot 98887+\cdot 0036/8 = 1\cdot 98932$.

The error in the approximation c_2 is less than $c_2-c_1 = \cdot 00045$, and so the root lies between $1\cdot 98887$ and $1\cdot 98977$, so that, correct to two places, its value is $1\cdot 99$.

13.31. Since $D_x(\sin x/x) = (\cos x/x-\sin x/x^2) = -1/\pi$ when $x = \pi$, it follows that the function $\sin x/x$ has a unique inverse near $x = \pi$ and so the equation $\sin x/x = \theta$ a unique solution near $x = \pi$.

Differentiating $\sin x = x\theta$ with respect to θ we find in turn

$$(\cos x-\theta)dx/d\theta = x, \qquad -\sin x(dx/d\theta)^2+(\cos x-\theta)d^2x/d\theta^2 = 2dx/d\theta,$$
$$-\cos x(dx/d\theta)^3-3\sin x(dx/d\theta)d^2x/d\theta^2+(\cos x-\theta)d^3x/d\theta^3 = 3d^2x/d\theta^2;$$

hence when $x = \pi$ and $\theta = 0$ we have $dx/d\theta = -\pi$, $d^2x/d\theta^2 = 2\pi$, $d^3x/d\theta^3 = -\pi^3-6\pi$ and therefore, by Theorem 13.521,

$$x = \pi-\pi\theta+\pi\theta^2-\pi(1+\pi^2/6)\theta^3+\theta^3\alpha_\theta,$$

where $\alpha_\theta \to 0$ when $\theta \to 0$.

13.32. If $f(x) = \frac{3}{2}x-\sin x$, then $f'(x) = \frac{3}{2}-\cos x$, $f''(x) = \sin x$. Hence, $|f''(x)| \leqslant 1$. Take $x = \frac{1}{2}\pi$ as a first approximation; then

$$f(\tfrac{1}{2}\pi) = (\tfrac{3}{4}\pi-1) = \cdot 0472,$$

and if $k = \frac{2}{3}$ then $kf'(\frac{1}{2}\pi) = \frac{2}{3}$ and so $\cdot 0472 = f(\frac{1}{2}\pi) < 1/12k^2M = \cdot 05\dot{3}$. Then

$$c_1 = \tfrac{1}{2}\pi-\tfrac{2}{3}(\cdot 0472) = 1\cdot 5118, \qquad f(c_1) = 1\cdot 0079-\cdot 9983 = \cdot 0096,$$
$$c_2 = 1\cdot 5118-\tfrac{2}{3}(\cdot 0096) = 1\cdot 4998, \qquad f(c_2) = \cdot 9999-\cdot 9986 = \cdot 0013,$$
$$c_3 = 1\cdot 4998-\tfrac{2}{3}(\cdot 0013) = 1\cdot 4982.$$

The difference between c_2 and c_3 is $\cdot 0015$ and so the root lies between $1\cdot 4967$ and $1\cdot 4998$, so that its value correct to 2 decimal places is $1\cdot 50$.

13.4. Since $D_y(y^3+y) = 3y^2+1 = 1$ when $y = 0$, the function y^3+y has a unique inverse near $y = 0$. Thus $y^3+y = x$ has a unique solution near $y = 0, x = 0$.

Differentiating with respect to x, we find in turn

$$(3y^2+1)dy/dx = 1, \qquad 6y(dy/dx)^2+(3y^2+1)d^2y/dx^2 = 0,$$
$$6(dy/dx)^3+18y(dy/dx)(d^2y/dx^2)+(3y^2+1)d^3y/dx^3 = 0,$$

$36(dy/dx)^2(d^2y/dx^2)+18y(d^2y/dx^2)^2+24y(dy/dx)(d^3y/dx^3)+$
$$+(3y^2+1)\,d^4y/dx^4 = 0.$$
$90(dy/dx)(d^2y/dx^2)^2+60(dy/dx)^2(d^3y/dx^3)+60y(d^2y/dx^2)(d^3y/dx^3)+$
$$+30y(dy/dx)(d^4y/dx^4)+(3y^2+1)d^5y/dx^5 = 0,$$
whence, as $y = 0$; $dy/dx = 1$, $d^2y/dx^2 = 0$, $d^3y/dx^3 = -6$, $d^4y/dx^4 = 0$, $d^5y/dx^5 = 360$, and therefore $y = x-x^3+3x^5(1+\epsilon_x)$.

13.41. $D_y(\log y/y) = (1-\log y)/y^2 = 1$ when $y = 1$ and so $\log y/y = x$, i.e. $y = e^{xy}$, has a unique solution for y in terms of x, near $y = 1$.

From $y = e^{xy}$ we have $\log y = xy$,
$$dy/dx = y^2+xy\,dy/dx, \qquad d^2y/dx^2 = 3y\,dy/dx+x(dy/dx)^2+xy\,d^2y/dx^2$$
and
$$d^3y/dx^3 = 4(dy/dx)^2+4y\,d^2y/dx^2+3x(dy/dx)(d^2y/dx^2)+xy\,d^3y/dx^3,$$
and therefore when $x = 0$ and $y = 1$ we have $dy/dx = 1$, $d^2y/dx^2 = 3$, $d^3y/dx^3 = 16$, whence the result follows.

13.5. On $y = ax^2+bx^3+cx^4$ at the point $x = 0$, we have $dy/dx = 0$, $d^2y/dx^2 = 2a$, $d^3y/dx^3 = 6b$, $d^4y/dx^4 = 24c$, $d^5y/dx^5 = 0$.

The general equation of a conic through the point $(0,0)$ is
$$Ax^2+2Hxy+By^2+2Gx+2Fy = 0,$$
which contains effectively four parameters.

On the conic, denoting successive derivatives by y', y'', and so on, we have
$$Ax+Hy+G+(Hx+By+F)y' = 0,$$
$$A+2Hy'+By'^2+(Hx+By+F)y'' = 0,$$
$$3Hy''+3By'y''+(Hx+By+F)y''' = 0,$$
$$4Hy'''+4By'y'''+3By''^2+(Hx+By+F)y^{\text{iv}} = 0.$$

For contact of the fourth order at $x = 0$, $y = 0$, we require $y' = 0$, $y'' = 2a$, $y''' = 6b$, $y^{\text{iv}} = 24c$, and therefore
$$G = 0, \quad A+2aF = 0, \quad 6Ha+6bF = 0, \quad 24Hb+12Ba^2+24Fc = 0,$$
whence
$$A/F = -2a, \qquad H/F = -b/a, \qquad B/F = 2(b^2-ac)/a^3$$
and the conic is $\qquad a^3y = a^4x^2+a^2bxy+(ac-b^2)y^2.$

Contact of the fifth order requires $y^{\text{v}} = 0$; but
$$5Hy^{\text{iv}}+5By'y^{\text{iv}}+10By''y'''+(Hx+By+F)y^{\text{v}} = 0$$
and therefore $Hc+Bab = 0$, whence $2b^3 = 3abc$, i.e. the conic has contact of the fifth order only if $b = 0$ or $2b^2 = 3ac$.

13.51. The general equation of a parabola through the origin is
$$(ax+by)^2 = 2px+2qy,$$
which contains effectively three parameters, and so on the parabola
$$(ax+by)(a+by') = p+qy', \qquad (ax+by)by''+(a+by')^2 = qy'',$$
$$(ax+by)by'''+(a+by')by''+2(a+by')by'' = qy'''.$$

If the parabola has contact of the third order at the origin with the curve $y = f(x)$ which touches the x-axis, then at the origin

$$y' = f'(0) = 0, \qquad y'' = f''(0), \qquad y''' = f'''(0);$$

but $f'(x) = \tan\psi$ and so

$$f''(x) = \sec^2\psi\frac{d\psi}{ds}\frac{ds}{dx} = \sec^3\psi/\rho \quad \text{and} \quad f'''(x) = 3\sec^5\psi\sin\psi/\rho^2 - \sec^4\psi\frac{d\rho}{ds}\Big/\rho^2,$$

whence $\qquad f''(0) = 1/\rho, \qquad f'''(0) = -\rho'/\rho^2.$

Thus from the conditions on y', y'', y''' we have

$p = 0$, $a^2 = q/\rho$, $3ab/\rho = -q\rho'/\rho^2$, whence $a = \sqrt{(q/\rho)}$, $b = -(\rho'/3)\sqrt{(q/\rho)}$ and the equation of the parabola is $(3x - y\rho')^2 = 18\rho y$.

For contact of the fourth order we require $4aby''' + 3b^2y''^2 = qy^{\text{iv}}$, where $y'' = f''(0) = 1/\rho$, $y''' = f'''(0) = -\rho'/\rho^2$, $y^{\text{iv}} = f^{\text{iv}}(0) = (3 + 2\rho'^2 - \rho\rho'')/\rho^3$, whence $5\rho'^2 = 9 + 6\rho'^2 - 3\rho\rho''$. Thus the condition for contact of the fourth order is $\rho'^2 + 9 = 3\rho\rho''$.

13.6. We have

$dx/ds = \cos\psi$, $d^2x/ds^2 = -\sin\psi/\rho$, $d^3x/ds^3 = -\cos\psi/\rho^2 + (\sin\psi/\rho^2)(d\rho/ds)$, and

$dy/ds = \sin\psi$, $d^2y/ds^2 = \cos\psi/\rho$, $d^3y/ds^3 = -\sin\psi/\rho^2 - (\cos\psi/\rho^2)(d\rho/ds)$;

at the origin $x = 0$, $y = 0$ we have $\psi = 0$ and so $dx/ds = 1$, $d^2x/ds^2 = 0$, $d^3x/ds^3 = -1/\rho^2$, and $dy/ds = 0$, $d^2y/ds^2 = 1/\rho$, $d^3y/ds^3 = -\rho'/\rho^2$, whence

$$x = s - s^3/3!\,\rho^2 + \ldots \quad \text{and} \quad y = s^2/2!\,\rho - \rho' s^3/3!\,\rho^2 + \ldots.$$

At the origin $d^4x/ds^4 = \rho'(\rho + 2\rho')/\rho^3$, $d^4y/ds^4 = \{2\rho'^2 - \rho(\rho' + \rho'')\}/\rho^3$ and so to the fourth power of s the expansions are

$$x = s - s^3/3!\,\rho^2 + \rho'(\rho + 2\rho')s^4/4!\,\rho^3,$$
$$y = s^2/2!\,\rho - \rho' s^3/3!\,\rho^2 + \{2\rho'^2 - \rho(\rho' + \rho'')\}s^4/4!\,\rho^3.$$

13.601. (i) If $s_n(x) = \sqrt{n}\,x(1-x)^n$ then, when

$$x = 1, \qquad s_n(x) = 0,$$
$$0 \leqslant x \leqslant 1/n, \qquad s_n(x) \leqslant x\sqrt{n} \leqslant 1/\sqrt{n},$$
$$1/n \leqslant x < 1, \qquad s_n(x) = (x\sqrt{n})/(1-x)^{-n} < (x\sqrt{n})/nx = 1/\sqrt{n},$$

since $(1-x)^{-n} = 1 + nx + \text{positive terms}$.

Hence for all x in $(0, 1)$, $0 \leqslant s_n(x) \leqslant 1/\sqrt{n}$, which proves that $s_n(x)$ is interval convergent in $(0, 1)$, with limit zero.

(ii) If $s_n(x) = x^n(1-x)$ then, when

$$1 - 1/\sqrt{n} \leqslant x \leqslant 1, \qquad s_n(x) \leqslant 1 - x = 1/\sqrt{n},$$
$$0 \leqslant x \leqslant 1 - 1/\sqrt{n}, \qquad s_n(x) \leqslant (1 - 1/\sqrt{n})^n = 1/(1 - 1/\sqrt{n})^{-n} < 1/\sqrt{n},$$

since $(1 - 1/\sqrt{n})^{-n} = 1 + \sqrt{n} + \text{positive terms}$;

thus for all x in $(0, 1)$, $0 \leqslant s_n(x) \leqslant 1/\sqrt{n}$, so that $s_n(x)$ is interval convergent with limit zero.

(iii) If $s_n(x) = nx(1-x)^n$, then $s_n(0) = s_n(1) = 0$ and when $0 < x < 1$,

$$s_n(x) = nx/(1-x)^{-n} < 2nx/n(n+1)x^2 \to 0$$

for a fixed x, so that $\lim s_n(x) = 0$ for any *fixed* x in $(0,1)$. But however great n may be chosen, $s_n(x)$ is not small for all values of x in $(0,1)$, since $s_n(1/n) = (1-1/n)^n \to 1/e$, so that $s_n(1/n) > \frac{1}{3}$ however great n may be.

13.61. We have $D^n e^{x\cos\alpha}\sin(x\sin\alpha) = e^{x\cos\alpha}\sin(x\sin\alpha+n\alpha)$, and so if $f(x) = e^{x\cos\alpha}\sin(x\sin\alpha)$, then $|f^n(x)| \leqslant e^{x\cos\alpha} \leqslant e^{a|\cos\alpha|}$, if $|x| \leqslant a$, and

$$f^n(0) = \sin n\alpha.$$

Since $f^n(x)$ is bounded, $f(x)$ equals its Taylor series and so

$$f(x) = \sum \frac{x^n}{n!}\sin n\alpha.$$

Similarly
$$e^{x\cos\alpha}\cos(x\sin\alpha) = \sum \frac{x^n}{n!}\cos n\alpha.$$

13.62. Since $\sum u_n$ converges we can find n_k so that $u_n+u_{n+1}+\ldots+u_N < \dfrac{1}{k}$ when $N \geqslant n \geqslant n_k$, and so

$$|a_n(x)+a_{n+1}(x)+\ldots+a_N(x)| \leqslant |a_n(x)|+|a_{n+1}(x)|+\ldots+|a_N(x)|$$
$$\leqslant u_n+u_{n+1}+\ldots+u_N(x) < \frac{1}{k}, \quad N \geqslant n \geqslant n_k,\ a \leqslant x \leqslant b,$$

where n_k is independent of x (depending only upon the series $\sum u_n$), which proves that $\sum a_n(x)$ is interval convergent in (a,b).

13.63.
$$\int\limits_0^c \frac{\phi(x)}{x^{\frac{1}{2}}(c-x)^{\frac{1}{2}}}\,dx = 2\int\limits_0^{\frac{1}{2}\pi} \phi(c\sin^2\theta)\,d\theta, \quad x = c\sin^2\theta;$$

but $\phi(x) = \sum\limits_{r=0}^{n} \phi^r(0)x^r/r!+x^n\epsilon(x)$, where $\epsilon(x) \to 0$ as $x \to 0$ and therefore

$$\int\limits_0^c \frac{\phi(x)}{x^{\frac{1}{2}}(c-x)^{\frac{1}{2}}}\,dx = 2\sum_{r=0}^{n}\phi^r(0)\left(\int\limits_0^{\frac{1}{2}\pi}\sin^{2r}\theta\,d\theta\right)c^r/r!+2c^n\int\limits_0^{\frac{1}{2}\pi}\epsilon(c\sin^2\theta)\sin^{2n}\theta\,d\theta.$$

Moreover.

$$\int\limits_0^{\frac{1}{2}\pi}\epsilon(c\sin^2\theta)\sin^{2n}\theta\,d\theta = \epsilon(c\sin^2\alpha)\int\limits_0^{\frac{1}{2}\pi}\sin^{2n}\theta\,d\theta, \quad 0 < \alpha < \tfrac{1}{2}\pi,$$
$$= \delta_c \quad \text{(say)},$$

so that $\delta_c \to 0$ as $c \to 0$.

13.64. Since the contact is of the nth order, and

$$f(a+h) = f(a)+hf'(a)+\frac{h^2}{2!}f''(a)+\ldots+\frac{h^n}{n!}f^n(a)+\frac{h^{n+1}}{(n+1)!}(f^{n+1}(a)+\alpha_h),$$

$$g(a+h) = g(a)+hg'(a)+\frac{h^2}{2!}g''(a)+\ldots+\frac{h^n}{n!}g^n(a)+\frac{h^{n+1}}{(n+1)!}(g^{n+1}(a)+\beta_h),$$

where $\alpha_h \to 0$, $\beta_h \to 0$ when $h \to 0$ and $f^{n+1}(a) \neq g^{n+1}(a)$, therefore

$$f(a+h)-g(a+h) = \frac{h^{n+1}}{(n+1)!}\{f^{n+1}(a)-g^{n+1}(a)+\gamma_h\}, \quad \text{where } \gamma_h = \alpha_h - \beta_h.$$

Thus for sufficiently small values of $|h|$, $\dfrac{f(a+h)-g(a+h)}{h^{n+1}}$ is of constant sign, since $f^{n+1}(a)-g^{n+1}(a)+\gamma_h$ is of constant sign when h is small enough to make $|\gamma_h| < |f^{n+1}(a)-g^{n+1}(a)|$. Accordingly $f(a+h)-g(a+h)$ has the same sign as h^{n+1} and so $f(a+h)-g(a+h)$ changes sign, and the curves cross, when n is even, and does not change sign, and the curves do not cross, when n is odd.

13.7. Since $\sum v_r(x)$ converges in $(0, \infty]$, we can find n_k so that for any $N > n_k$ and for all positive x, $\left|\sum_{n_k+1}^{N} v_r(x)\right| < 1/k$, and so

$$\left|\sum_{n_k+1}^{N} \quad \right| \quad \left|\sum_{n_k+1}^{N} \lim_{x\to\infty} v_r(x)\right| = \lim \left|\sum_{n_k+1} v_r(x)\right| \quad 1/k,$$

which proves that $\sum w_r$ converges; incidentally we observe

$$\left|\sum_{r\geqslant n_k+1} v_r(x)\right| \quad \left|\lim_{N\to\infty}\left\{\sum_{n_k+1}^{N} v_r(x)\right\}\right| \quad 1/k,$$

and similarly $\left|\sum_{r>n_k} w_r\right| \quad 1/k.$

Since

$$\left|\sum_{1}^{p_n} v_r(x) - \sum w_r\right| \leqslant \left|\sum_{1}^{n_k} (v_r(x)-w_r)\right| + \left|\sum_{n_k+1}^{p_n} v_r(x)\right| + \left|\sum_{r\geqslant n_k+1} w_r\right|,$$

and since we can choose N_k so that for $n \geqslant N_k$

$$|v_r(n)-w_r| < 1/kn_k, \quad r = 0, 1, 2,\ldots, n_k,$$

therefore, when $n \geqslant N_k$ and $p_n > n_k$

$$\left|\sum^{p_n} v_r(n) - \sum w_r\right| \leqslant 3/k, \quad \text{i.e.} \quad \sum^{p_n} v_r(n) \to \sum w_r.$$

13.71. Since $\log(1+x)^{1/x} \to 1$ as $x \to 0$, we can find x_0 so that

$$\left|\frac{\log(1+x)}{x}\right| < 2, \quad |x| \leqslant x_0.$$

As $\sum |v_r(x)|$ converges in $(0, \infty]$, we can find n_0 so that $|v_r(x)| < x_0$ when $r \geqslant n_0$. Hence

$$|\log\{1+v_r(x)\}| < 2|v_r(x)|, \quad r \geqslant n_0.$$

Thus $\sum \log\{1+v_r(x)\}$ converges in the interval $(0, \infty]$, and

$$\lim_{x\to\infty} \log\{1+v_r(x)\} = \log(1+w_r).$$

Hence, by 13.7, $\sum \log(1+w_r)$ converges, and

$$\sum_{r=1}^{p_n} \log\{1+v_r(n)\} \to \sum \log(1+w_r),$$

and therefore
$$\prod_{r=1}^{p_n} \{1+v_r(n)\} \to \prod_{r\geqslant 1} (1+w_r),$$

for $$e^{\sum_1^{p_n}\log\{1+v_r(n)\}} = \prod_1^{p_n} e^{\log\{(1+v_r(n))\}} = \prod_1^{p_n}\{1+v_r(n)\}$$
and
$$e^{\sum\log(1+w_r)} = e^{\lim\sum_1^n\log(1+w_r)} = \lim\left\{e^{\sum_1^n\log(1+w_r)}\right\}, \quad \text{since } e^x \text{ is continuous,}$$
$$= \lim\prod_1^n(1+w_r) = \prod_{r\geqslant 1}(1+w_r).$$

13.72. It follows from Example 5.2 that if n is odd $\sin n\theta$ is a polynomial $p(\sin\theta)$ of the nth degree in $\sin\theta$; since $\sin n\theta = 0$ when $\theta = \pm r\pi/n$, $r = 0, 1, 2,\ldots, (n-1)/2$, the roots of $p(t) = 0$ are $t = \pm\sin r\pi/n$, $r = 0, 1,\ldots, (n-1)/2$.

Writing t for $\sin\theta$ we have
$$\sin n\theta = At(t^2-\sin^2\pi/n)(t^2-\sin^2 2\pi/n)(t^2-\sin^2 3\pi/n)\ldots[t^2-\sin^2\{(n-1)/2\}\pi/n]$$
i.e.
$$\sin n\theta/\sin\theta$$
$$= B(1-\sin^2\theta/\sin^2\pi/n)(1-\sin^2\theta/\sin^2 2\pi/n)\ldots[1-\sin^2\theta/\sin^2\{(n-1)/2\}\pi/n].$$
Since $\lim_{\theta\to 0}\sin n\theta/\sin\theta = \lim_{\theta\to 0} n\cos n\theta/\cos\theta = n$, therefore $B = n$.

Write $n\theta = \phi$, then
$$(\sin\phi/\phi)/\{\sin(\phi/n)/(\phi/n)\} = \prod_{r=1}^{\frac{1}{2}(n-1)}\{1-\sin^2(\phi/n)/\sin^2(r\pi/n)\}.$$
Let $v_r(x) = -\sin^2(\phi/x)/\sin^2(r\pi/x)$ for $x \geqslant 2r$ and $v_r(x) = 0$ for $x < 2r$, then since $|\sin\alpha/\alpha| \leqslant 1$ for any α and $\sin\alpha/\alpha \geqslant 2/\pi$ provided $0 < \alpha \leqslant \frac{1}{2}\pi$, we have, if $x \geqslant 2r$ (and therefore $r\pi/x \leqslant \frac{1}{2}\pi$),
$$|v_r(x)| = [\{(\sin\phi/x)/(\phi/x)\}/\{(\sin r\pi/x)/(r\pi/x)\}]^2(\phi^2/r^2\pi^2) < (\tfrac{1}{2}\pi)^2(\phi^2/r^2\pi^2)$$
$$= \phi^2/4r^2$$
and if $x < 2r$, $|v_r(x)| = 0 < \phi^2/4r^2$, and so $\sum_p^q|v_r(x)| < (\phi^2/4)\sum_p^q 1/r^2$ for any x; but $\sum 1/r^2$ converges and therefore $\sum|v_r(x)|$ converges in $(0,\infty)$ for any fixed ϕ.

Furthermore, since $(\sin\alpha/x)/(\alpha/x) \to 1$ as $x\to\infty$, for any α, therefore $v_r(x) \to -\phi^2/r^2\pi^2$ as $x\to\infty$. Thus $v_r(x)$ satisfies the conditions of Example 13.71 and therefore
$$\prod_{r=1}^{\frac{1}{2}(n-1)}\{1-\sin^2(\phi/n)/\sin^2(r\pi/n)\} = \prod_{r=1}^{\frac{1}{2}(n-1)}\{1+v_r(n)\} \to \prod_{r\geqslant 1}(1-\phi^2/r^2\pi^2);$$
but $(\sin\phi/\phi)/\{(\sin\phi/n)/(\phi/n)\} \to \sin\phi/\phi$, whence
$$\sin\phi = \phi\prod_{r\geqslant 1}(1-\phi^2/r^2\pi^2) \quad \text{for any } \phi.$$

13.8. If $\phi(h) = f(a)+r\{f(a+h)-f(a)\}-f(a+rh)$, so that $\phi(0) = 0$, then $\phi'(h) = rf'(a+h)-rf'(a+rh) = r(1-r)hf''\{\alpha(h)\}$, $a+h > \alpha(h) > a+rh$.
Hence by the Cauchy formula
$$\frac{\phi(h)}{h} = \frac{\phi'(c)}{2c} = \frac{r(1-r)cf''\{\alpha(c)\}}{2c}, \quad 0 < c < h,$$

whence $\phi(h) = \frac{1}{6}r(1-r)h^2f''\{\alpha(c)\}$; since $a < a+rc < \alpha(c) < a+c < a+h$, therefore $\alpha(c)$ is of the form $a+\theta h$, $0 < \theta < 1$. (Alternatively deduce from Theorem 12.541 writing $a = y$, $a+h = z$, $a+rh = x$).

Take $f(x) = \sin x$, $a = \frac{1}{6}\pi$, $h = \frac{1}{12}\pi$ and $r = \frac{1}{3}$. Then an approximation to

$$\sin\frac{7\pi}{36} = \sin\left(\frac{\pi}{6}+\frac{1}{3}\cdot\frac{\pi}{12}\right) \quad \text{is} \quad \sin\frac{\pi}{6}+\frac{1}{3}\left(\sin\frac{\pi}{4}-\sin\frac{\pi}{6}\right) : \frac{2+\sqrt{2}}{6}$$

with an error
$$\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{2}{3}\cdot\frac{\pi^2}{144}\sin\frac{\pi}{12}(\theta+2) < \frac{\pi^2\sqrt{2}}{2592}.$$

13.81. Let $a-x = y$, so that $\delta \leqslant y \leqslant a$; then if $y > 1$ and if $y = 1+d$, so that $0 < d \leqslant a-1$, it follows from $\left(1+\frac{d}{n}\right)^n > 1+d$ that

$$1 < y^{1/n} < 1+\frac{d}{n} \leqslant 1+\frac{a-1}{n},$$

whence $|y^{1/n}-1| \leqslant \dfrac{a-1}{n} \leqslant \delta$, if $n \geqslant (a-1)/\delta$. If $y = 1$, $|y^{1/n}-1| = 0 < \delta$, for all n. If $y < 1$, and if $z = 1/y$, then $1/\delta \geqslant z > 1$ and

$$|y^{1/n}-1| = \left|\frac{1-z^{1/n}}{z^{1/n}}\right| < |1-z^{1/n}| \leqslant \delta, \quad \text{if } n \geqslant (\delta^{-1}-1)/\delta.$$

Thus in every case, if $0 \leqslant x \leqslant a-\delta$ then $|(a-x)^{1/n}-1| < \delta$ for $n \geqslant n_\delta$. Similarly, if $|f(x)| \leqslant M$, then for $0 \leqslant x \leqslant a-\epsilon_\delta$,

$$|(a-x)^{1/n}f(x)-f(x)| < \frac{\delta}{M}|f(x)| < \delta, \quad \text{for } n \geqslant \nu_\delta,$$

and for $0 \leqslant a-x \leqslant \epsilon_\delta < 1$,

$$|(a-x)^{1/n}f(x)-f(x)| = |f(x)| < \delta, \quad \text{for all } n,$$

so that $\qquad |f_n(x)-f(x)| < \delta, \quad n \geqslant \nu_\delta$, and all x in $(0,a)$.

Furthermore $\qquad f_n'(x) = (a-x)^{1/n}f'(x)-\dfrac{1}{n}(a-x)^{1/n}\dfrac{f(x)}{(a-x)}$,

and by the mean-value theorem, since $f(a) = 0$, $-f(x) = (a-x)f'\{c(x)\}$, where $x < c(x) < a$, so that

$$f_n'(x) = (a-x)^{1/n}f'(x)+\frac{1}{n}(a-x)^{1/n}f'\{c(x)\};$$

hence, if $|f'(x)| < K$, then when $0 \leqslant x \leqslant a-\eta_\delta$, $(\eta_\delta < 1)$

$$\begin{aligned}|f_n'(x)-f'(x)| &\leqslant |(a-x)^{1/n}-1|K+\frac{1}{n}(a+1)K \\ &< \tfrac{1}{2}\delta+\frac{a+1}{n}K, \quad n \geqslant \mu_\delta, \\ &< \delta \quad \text{if} \quad n > \max\{2(a+1)K/\delta, \mu_\delta\}.\end{aligned}$$

If $0 \leqslant a-x \leqslant \eta_\delta$, $|f_n'(x)-f'(x)| \leqslant |f'(x)|+|f'\{c(x)\}|/n < \delta$ by continuity.

13.82. $\dfrac{dy}{dx} = -e^{-x}\left(1+\dfrac{x}{n}\right)^n+e^{-x}\left(1+\dfrac{x}{n}\right)^{n-1}$ and so $(x+n)\dfrac{dy}{dx}+xy = 0$; hence, in turn,

$$(x+n)\frac{d^2y}{dx^2}+(1+x)\frac{dy}{dx}+y = 0, \qquad (x+n)\frac{d^2y}{dx^2}-(1+x)\frac{xy}{x+n}+y = 0,$$

$$\frac{d^3y}{dx^3} = (2n+3nx-x^2)y/(x+n)^2,$$

whence the result follows by Taylor's theorem. Furthermore

$$\left|n\left[1-e^{-x}\left(1+\frac{x}{n}\right)^n\right]-\frac{x^2}{2}\right| = \frac{x^3}{3!}\frac{n(2n+3nc-c^2)}{(c+n)^3}e^{-c}\left(1+\frac{c}{n}\right)^n$$

$$< \frac{x^3}{3!}\left(\frac{2}{n^2}+\frac{3x}{n^2}+\frac{x^2}{n^3}\right)\left(1+\frac{x}{n}\right)^n$$

$$< \frac{x^3}{3!}e^x\left(\frac{2}{n^2}+\frac{3x}{n^2}+\frac{x^2}{n^3}\right) \to 0.$$

13.83. (i) Let $x > X$; by the mean-value theorem we can find c, $X < c < x$, such that

$$\frac{f(x)-f(X)}{x-X} = f'(c); \quad \text{therefore} \quad \frac{f(x)}{x} = \frac{f(X)}{x}+\left(1-\frac{X}{x}\right)f'(c).$$

Choose X so great that $l-\frac{1}{k} < f'(c) < l+\frac{1}{k}$, $k > 1$, and keeping X fixed choose x so great that

$$\left|\frac{f(X)}{x}\right| < \frac{1}{k}, \qquad \left|\frac{X}{x}\right| < \frac{1}{k},$$

then
$$\frac{f(x)}{x} < \frac{1}{k}+\left(1+\frac{1}{k}\right)\left(l+\frac{1}{k}\right) < l+\frac{1}{k}(l+3),$$

and
$$\frac{f(x)}{x} > -\frac{1}{k}+\left(1-\frac{1}{k}\right)\left(l-\frac{1}{k}\right) > l-\frac{1}{k}(l+2)$$

whence
$$\frac{f(x)}{x} \to l, \quad \text{as} \quad x \to \infty.$$

(ii) $\phi'(x) = \dfrac{1}{\theta'(x)} \to \dfrac{1}{l} > 0$, therefore $\dfrac{\phi(x)}{x} \to \dfrac{1}{l}$; hence, since $\dfrac{\theta(x)}{x} \to l$, we have

$$\frac{\psi(x)}{x^2} = \frac{\theta(x)}{x}\cdot\frac{\phi(x)}{x} \to l\cdot\frac{1}{l} = 1.$$

13.9. We observe that if q and s are both positive then $\dfrac{p+r}{q+s}-\dfrac{p}{q}$ and $\dfrac{p+r}{q+s}-\dfrac{r}{s}$ have opposite signs, so that $\dfrac{p+r}{q+s}$ lies between $\dfrac{p}{q}$ and $\dfrac{r}{s}$.

Since both b_n-c and $c-a_n$ are positive and

$$\frac{f(b_n)-f(a_n)}{b_n-a_n} = \frac{\{f(b_n)-f(c)\}+\{f(c)-f(a_n)\}}{(b_n-c)+(c-a_n)}$$

therefore $\dfrac{f(b_n)-f(a_n)}{b_n-a_n}$ lies between $\dfrac{f(b_n)-f(c)}{b_n-c}$ and $\dfrac{f(c)-f(a_n)}{c-a_n}$, each of which has the limit ϕ (for $b_n-c < b_n-a_n \to 0$, $c-a_n < b_n-a_n \to 0$). Hence

$$\frac{f(b_n)-f(a_n)}{b_n-a_n} \to \phi.$$

13.91. Using Example 13.9 and the fact that the point-derivative $\phi(x)$ is given to be continuous the proof proceeds exactly as in §§ 3.61, 12.21.

13.92. By Example 13.91, given any x, X in (a,b) we can find x^* in $[x,X]$ such that $\{f(X)-f(x)\}/(X-x) = \phi(x^*)$; since $\phi(x)$ is continuous in (a,b) we can find p_n independent of x, X such that $\phi(X)-\phi(x) = 0(n)$ for any x, X in (a,b) satisfying $X-x = 0(p_n)$. Hence if $X-x = 0(p_n)$, then $x^*-x = 0(p_n)$ and therefore $\{f(X)-f(x)\}/(X-x) = \phi(x)+0(n)$, which proves that $f(x)$ is interval-differentiable in (a,b) with derivative $\phi(x)$.

XIV

14.3. Use Example 9.61 and integrate from 0 to 2π.

14.41. Use Example 14.4.

14.7. If $y = \sqrt{x}$ then

$$\int_1^{x_{n+1}} \frac{1}{y}\,dy = \frac{1}{2}\int_1^{x_n} \frac{1}{x}\,dx$$

and so

$$2^n \int_1^{x_n} \frac{1}{y}\,dy = \int_1^{x_0} \frac{1}{x}\,dx.$$

Since $x_{n+1} = \sqrt{x_n}$ therefore $\dfrac{x_{n+1}-1}{x_n-1} = \dfrac{1}{x_{n+1}+1}$; but if $x_0 > 1$ then $x_n > 1$ for all n and so $\dfrac{x_{n+1}-1}{x_n-1} < \dfrac{1}{2}$ and therefore $\dfrac{x_n-1}{x_0-1} < \dfrac{1}{2^n}$ so that $x_n \to 1$. If $x_0 < 1$ then $1/x_n > 1$ so that, by the foregoing proof, $1/x_n \to 1$, that is, $x_n \to 1$. Finally, if $x_0 = 1$, then all $x_n = 1$.

Next we observe that

$$\lim_{t\to 1} \frac{\int_1^t (1/x)\,dx}{t-1} = \lim_{t\to 1} 1/t = 1$$

and so

$$\frac{2^n \int_1^{x_n} (1/x)\,dx}{2^n(x_n-1)} = \frac{\int_1^{x_n} (1/x)\,dx}{x_n-1} \to 1,$$

which completes the proof.

14.71. Since $x_{n+1} = x_n/\{1+\sqrt{(1+x_n^2)}\}$ therefore $|x_{n+1}| < \frac{1}{2}|x_n|$ and so

$$|x_n| < \frac{1}{2^n}|x_0| \to 0.$$

Under the transformation $y = x/\{1+\sqrt{(1+x^2)}\} = \dfrac{\sqrt{(1+x^2)}-1}{x}$ we have

$$y+\frac{1}{y} = \frac{2\sqrt{(1+x^2)}}{x}, \qquad y-\frac{1}{y} = -\frac{2}{x},$$

and so

$$\left(y+\frac{1}{y}\right)\frac{1}{y}\frac{dy}{dx} = \left(1+\frac{1}{y^2}\right)\frac{dy}{dx} = \frac{2}{x^2},$$

whence $$\frac{1}{1+y^2}\frac{dy}{dx} = \frac{1}{y+(1/y)}\cdot\frac{1}{y}\frac{dy}{dx} = \frac{2}{x^3}\cdot\frac{x^3}{4(1+x^2)} = \frac{1}{2}\,\frac{1}{1+x^2}.$$

Hence $$\int_0^{x_{n+1}} \frac{1}{1+y^2}\,dy = \frac{1}{2}\int_0^{x_n} \frac{1}{1+x^2}\,dx$$

and therefore $$2^n \int_0^{x_n} \frac{1}{1+y^2}\,dy = \int_0^{x_0} \frac{1}{1+x^2}\,dx.$$

Since $$\lim_{t\to 0} \frac{\int_0^t \{1/(1+x^2)\}\,dx}{t} = \lim_{t\to 0}\frac{1}{1+t^2} = 1$$

therefore $$\frac{2^n \int_0^{x_n} \{1/(1+y^2)\}\,dy}{2^n x_n} = \frac{\int_0^{x_n} \{1/(1+y^2)\}\,dy}{x_n} \to 1$$

and so $2^n x_n \to \int_0^{x_0} \frac{1}{1+x^2}\,dx = \arctan x_0$.

14.8. By the mean-value theorem we can find c and c^* in (a,b) such that

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(c) \quad\text{and}\quad \frac{1}{b-a}\int_a^b \log f(x)\,dx = \log f(c^*)$$

and therefore $$A^f_{a,b} = f(c), \qquad G^f_{a,b} = f(c^*).$$

14.81. $$\int_a^\infty f(x,Y)\,dx - \int_a^\infty f(x,y)\,dx$$

$$= \int_a^n [f(x,Y)-f(x,y)]\,dx + \int_n^\infty f(x,Y)\,dx - \int_n^\infty f(x,y)\,dx.$$

By interval convergence

$$\left|\int_n^\infty f(x,Y)\,dx\right| < \frac{1}{k}, \qquad \left|\int_n^\infty f(x,y)\,dx\right| < \frac{1}{k}, \quad n > n_k.$$

By continuity,

$$|f(x,Y)-f(x,y)| < \frac{1}{kn_k} \quad\text{if}\quad |Y-y| < \frac{1}{p_k},$$

whence $$\left|\int_a^n \{f(x,Y)-f(x,y)\}\,dx\right| < \frac{1}{k}, \quad n = n_k,$$

which completes the proof.

14.82. If $\sum \int_a^\infty u_n(x)\,dx = S$ and $\int_a^\infty \sum u_n(x)\,dx = S'$,

then $$\int_a^n \sum u_n(x)\,dx = \sum \int_a^n u_n(x)\,dx \leqslant S \quad \text{for all } n,$$
and therefore $S' \leqslant S$.

Furthermore
$$\sum_1^N \int_a^\infty u_n(x)\,dx = \int_a^\infty \sum_1^N u_n(x)\,dx \leqslant \int_a^\infty \sum_1^\infty u_n(x)\,dx = S', \quad \text{for all } N,$$
and therefore $S \leqslant S'$, whence $S = S'$.

If $u_n(x)$ may have any sign but the conditions of the theorem are satisfied both for $u_n(x)$ and $|u_n(x)|$ then by addition and subtraction respectively, the conditions are satisfied for the *positive* functions $|u_n(x)| \pm u_n(x)$, whence by the first part
$$\sum \int_a^\infty \{|u_n(x)|+u_n(x)\}\,dx = \int_a^\infty \sum \{|u_n(x)|+u_n(x)\}\,dx$$
and $$\sum \int_a^\infty \{|u_n(x)|-u_n(x)\}\,dx = \int_a^\infty \sum \{|u_n(x)|-u_n(x)\}\,dx,$$
whence by subtraction
$$\sum \int_a^\infty u_n(x)\,dx = \int_a^\infty \sum u_n(x)\,dx.$$

14.9. Let q be greater than p, and let $a_r^p = a_s^q$, $a_{r+1}^p = a_t^q$, then
$$S_p = \sum_r f(a_r^p)(a_{r+1}^p - a_r^p) = \sum_r f(a_r^p)\{(a_{s+1}^q - a_s^q)+(a_{s+2}^q - a_{s+1}^q)+\dots+(a_t^q - a_{t-1}^q)\},$$
and
$$\begin{aligned} S_q &= \sum_\rho f(a_\rho^q)(a_{\rho+1}^q - a_\rho^q) \\ &= \sum \{f(a_s^q)(a_{s+1}^q - a_s^q) + f(a_{s+1}^q)(a_{s+2}^q - a_{s+1}^q) + \dots + f(a_{t-1}^q)(a_t^q - a_{t-1}^q)\} \end{aligned}$$
and therefore
$$S_q - S_p = \sum_r \left\{ \sum_{\rho=s}^{t-1} \{f(a_\rho^q) - f(a_s^q)\}(a_{\rho+1}^q - a_\rho^q) \right\} = 0(p) \sum_\rho (a_{\rho+1}^q - a_\rho^q) = (b-a).0(p),$$
which proves that S_p converges.

14.91. Exactly as in Theorem 14 we prove
$$\sum_r f(a_r^p)(a_{r+1}^p - a_r^p) - \sum_r f(b_r^p)(b_{r+1}^p - b_r^p) = (b-a).0(p-1).$$
Thus $S_p^1 - S_p^2 \to 0$; hence if l_1, l_2 are the limits of S_p^1, S_p^2 (convergent by 14.9), then $l_1 - l_2 = (l_1 - S_p^1) - (l_2 - S_p^2) + (S_p^1 - S_p^2) \to 0$, and so $l_1 = l_2$.

This proves that the integral of $f(x)$ is independent of the particular p-chain on which it is formed.

14.92. Let $a = a_0^p, a_1^p, a_2^p, \dots, a_i^p = c, a_{i+1}^p, \dots, a_j^p = b$ be a p-chain for the semi-continuous function $f(x)$. Then
$$\sum_{r=0}^{i-1} f(a_r^p)(a_{r+1}^p - a_r^p) \to \int_a^c f(x)\,dx, \qquad \sum_{r=i}^{j-1} f(a_r^p)(a_{r+1}^p - a_r^p) \to \int_c^b f(x)\,dx,$$

and $$\sum_{r=j}^{s-1} f(a_r^p)(a_{r+1}^p - a_r^p) \underset{p}{\to} \int_c^b f(x)\,dx,$$

so that $$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx.$$

14.93. Let $a = a_0^p, a_1^p, a_2^p, \ldots, a_{s_p}^p = b$ be a p-chain for all three functions $f(x)$, $g(x)$, and $f(x)+g(x)$, then

$$\sum_r f(a_r^p)(a_{r+1}^p - a_r^p) \to \int_a^b f(x)\,dx, \qquad \sum_r g(a_r^p)(a_{r+1}^p - a_r^p) \to \int_a^b g(x)\,dx,$$

and $$\sum_r \{f(a_r^p)+g(a_r^p)\}(a_{r+1}^p - a_r^p) \to \int_a^b \{f(x)+g(x)\}\,dx$$

so that $$\int_a^b f(x)\,dx + \int_a^b g(x)\,dx = \int_a^b \{f(x)+g(x)\}\,dx.$$

14.94. If $m \leqslant f(x) \leqslant M$ in $(a,b]$, then $S_p = \sum_r f(a_r^p)(a_{r+1}^p - a_r^p)$ lies between $m\sum_r (a_{r+1}^p - a_r^p)$ and $M\sum_r (a_{r+1}^p - a_r^p)$, i.e. between $(b-a)m$ and $(b-a)M$, and therefore $(b-a)m \leqslant \lim S_p \leqslant (b-a)M$.

14.941. $$\left|\int_a^T f(x)\,dx - \int_a^t f(x)\,dx\right| = \left|\int_a^t f(x)\,dx + \int_t^T f(x)\,dx - \int_a^t f(x)\,dx\right|$$
$$= \left|\int_t^T f(x)\,dx\right| \leqslant M(T-t),$$

where M is a bound of $|f(x)|$, which proves that $\int_a^t f(x)\,dx$ is continuous.

14.95. We prove first that if $a = a_0^p, a_1^p, a_2^p, \ldots, a_{s_p}^p = b$ is a p-chain for $f(x)$ then

$$\sum_r f(a_r^p) \int_{a_r^p}^{a_{r+1}^p} g(x)\,dx \to \int_a^b f(x)g(x)\,dx.$$

For

$$\left|\int_a^b f(x)g(x)\,dx - \sum_r f(a_r^p)\int_{a_r^p}^{a_{r+1}^p} g(x)\,dx\right| = \left|\sum_r \int_{a_r^p}^{a_{r+1}^p} \{f(x)-f(a_r^p)\}g(x)\,dx\right|$$

$\leqslant M\sum_r \int_{a_r^p}^{a_{r+1}^p} \{f(x)-f(a_r^p)\}\,dx,$ since $f(x)$ is non-decreasing, and $|g(x)| \leqslant M$,

$\leqslant M.\theta(p)\sum (a_{r+1}^p - a_r^p) = M(b-a).\theta(p).$

But

$$\Sigma_p = \sum_r f(a_r^p) \int_{a_r^p}^{a_{r+1}^p} g(x)\, dx = f(a)\left(\int_a^b - \int_{a_1^p}^b g(x)\, dx\right) + f(a_1^p)\left(\int_{a_1^p}^b - \int_{a_2^p}^b g(x)\, dx\right) + \ldots$$

$$= f(a) \int_a^b g(x)\, dx + \sum_{r=0}^{n_p-1} \{f(a_{r+1}^p) - f(a_r^p)\} \int_{a_{r+1}^p}^b g(x)\, dx;$$

since $f(a)$ and $f(a_{r+1}^p) - f(a_r^p)$ are all positive, and

$$f(a) + \sum_{r=0}^{n_p-1} \{f(a_{r+1}^p) - f(a_r^p)\} = f(b),$$

therefore if l and g are the least and greatest values of $\int_t^b g(x)\, dx$, where $a < t < b$, it follows that

$$lf(b) < \Sigma_p < gf(b).$$

But $\Sigma_p \to \int_a^b f(x)g(x)\, dx$, and so

$$l < \left(\int_a^b f(x)g(x)\, dx\right) \Big/ f(b) < g,$$

i.e. $\left(\int_a^b f(x)g(x)\, dx\right) \Big/ f(b)$ lies between two values of the *continuous* function $\int_t^b g(x)\, dx$, and therefore is itself a value of $\int_t^b g(x)\, dx$. Thus there is a point c in $[a, b]$ such that

$$\int_a^b f(x)g(x)\, dx = f(b) \int_c^b g(x)\, dx.$$

14.96. If for some k, $(k-1)\pi < |x| < k\pi$, and if $n > k$, then

$$n^2 - k^2 = (n-k)(n+k) > (n-k)^2$$

and so

$$\sum_n^N 2|x|/(r^2\pi^2 - x^2) < (2k/\pi) \sum_n^N 1/(r^2 - k^2) < (2k/\pi) \sum_n^N 1/(r-k)^2$$

$$= (2k/\pi) \sum_{n-k}^{N-k} 1/r^2 = 0(p), \quad n > n_p,$$

since $\sum 1/r^2$ converges; thus $\sum 2x/(r^2\pi^2 - x^2)$ is interval-convergent in the interval $[(k-1)\pi, k\pi]$, for any k.

But
$$D_x \log\left(1 - \frac{x^2}{r^2\pi^2}\right) = -\frac{2x}{r^2\pi^2 - x^2}$$

and, by Example 13.72,

$$\sum \log\left(1 - \frac{x^2}{r^2\pi^2}\right) = \log\sin x - \log x$$

and therefore, by Theorem 14.61,

$$D\{\log\sin x - \log x\} = -2x \sum \frac{1}{r^2\pi^2 - x^2},$$

i.e. $\qquad \cot x = \dfrac{1}{x} - 2x \displaystyle\sum_{1}^{\infty} \frac{1}{r^2\pi^2 - x^2}$ for any x not a multiple of π.

Since $\dfrac{2x}{x^2 - r^2\pi} = \dfrac{1}{x - r\pi} + \dfrac{1}{x + r\pi}$ the result may be written

$$\cot x = \sum_{r=-\infty}^{\infty} \frac{1}{x - r\pi}.$$

14.961. By Example 14.96

$$\tan \tfrac{1}{2}x = -\cot(\tfrac{1}{2}x - \tfrac{1}{2}\pi) = -\sum_{-\infty}^{\infty} \frac{1}{\frac{1}{2}x - (r+\frac{1}{2})\pi} = -\sum_{-\infty}^{\infty} \frac{2}{x - (2r+1)\pi}$$

and therefore

$$\frac{1}{\sin x} = \tfrac{1}{2}(\cot \tfrac{1}{2}x + \tan \tfrac{1}{2}x) = \sum_{-\infty}^{\infty} \frac{1}{x - 2r\pi} - \sum_{-\infty}^{\infty} \frac{1}{x - (2r+1)\pi} = \sum_{-\infty}^{\infty} \frac{(-1)^r}{x - r\pi}.$$

14.962. Since

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-x)^{n-1} + \frac{(-x)^n}{1+x} \quad \text{for any } n,$$

therefore

$$\int_0^1 \frac{x^{a-1}}{1+x}\,dx = \sum_0^{n-1} (-1)^r \int_0^1 x^{a+r-1}\,dx + (-1)^n \int_0^1 \frac{x^{a+n-1}}{1+x}$$

$$= \sum_0^{n-1} \frac{(-1)^r}{a+r} + \frac{(-1)^n}{1+\delta} \cdot \frac{1}{a+n}, \quad \text{where } 0 < \delta < 1, \quad \text{for any } n,$$

$$\to \sum_0^{\infty} \frac{(-1)^r}{a+r}, \text{ the series converging since } \frac{1}{a+n} \text{ is steadily decreasing to zero.}$$

Thus
$$\int_0^1 \frac{x^{a-1}}{1+x}\,dx = \sum_0^{\infty} \frac{(-1)^r}{a+r};$$

but
$$\int_{1/n}^1 \frac{x^{-a}}{1+x}\,dx = \int_1^n \frac{y^{a-1}}{1+y}\,dy, \quad x = y^{-1},$$

$$\to \int_1^{\infty} \frac{x^{a-1}}{1+x}\,dx,$$

and so

$$\int_1^{\infty} \frac{x^{a-1}}{1+x}\,dx = \int_0^1 \frac{x^{-a}}{1+x}\,dx = \sum_0^{\infty} \frac{(-1)^r}{1-a+r} = \sum_1^{\infty} \frac{(-1)^k}{a-k},$$

since $1-a$ is positive.

whence

$$\int_0^\infty \frac{x^{a-1}}{1+x}\,dx = \int_0^1 \frac{x^{a-1}}{1+x}\,dx + \int_1^\infty \frac{x^{a-1}}{1+x}\,dx$$

$$= \sum_0^\infty \frac{(-1)^r}{a+r} + \sum_1^\infty \frac{(-1)^k}{a-k} = \sum_0^{-\infty} \frac{(-1)^k}{a-k} + \sum_1^\infty \frac{(-1)^k}{a-k} = \sum_{-\infty}^\infty \frac{(-1)^k}{a-k}$$

$$= \pi \sum_{-\infty}^\infty \frac{(-1)^k}{a\pi - k\pi} = \frac{\pi}{\sin a\pi}.$$

XV

15. Taking x, y as the independent, and u, v as the dependent variables we find from the first equation, taking x and then y as the variable,

$$\cos u = v\frac{\partial u}{\partial x}, \qquad \sin u = v\frac{\partial u}{\partial y},$$

whence
$$v\frac{\partial^2 u}{\partial x\partial y}+\frac{\partial u}{\partial x}\frac{\partial v}{\partial y}+\sin u\frac{\partial u}{\partial y} = 0.$$

From the equation $x\sin u - y\cos u = v$, taking y as the variable, we find

$$-\cos u+\frac{\partial u}{\partial y} = \frac{\partial v}{\partial y};$$

thus
$$v^2\frac{\partial^2 u}{\partial x\partial y}+v\frac{\partial u}{\partial x}\left(\frac{\partial u}{\partial y}-\cos u\right)+\sin u.v\frac{\partial u}{\partial y} = 0,$$

i.e.
$$v^2\frac{\partial^2 u}{\partial x\partial y}+v\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}-\cos^2 u+\sin^2 u = 0.$$

15.01.
$$\frac{\partial H}{\partial x} = \frac{\partial H}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial H}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial H}{\partial u}+\frac{\partial H}{\partial v}$$

and so
$$\frac{\partial^2 H}{\partial x^2} = \left(\frac{\partial}{\partial u}+\frac{\partial}{\partial v}\right)^2 H;$$

similarly
$$\frac{\partial^2 H}{\partial y^2} = n^2\left(\frac{\partial}{\partial u}-\frac{\partial}{\partial v}\right)^2 H,$$

and so
$$n^2\frac{\partial^2 H}{\partial x^2} = \frac{\partial^2 H}{\partial y^2} \quad \text{becomes} \quad \frac{\partial^2 H}{\partial u\partial v} = 0;$$

therefore $\partial H/\partial v$ is independent of u, i.e. $\partial H/\partial v = \phi(v)$, and so $H-\int \phi(v)\,dv$ is independent of v, wherefore $H = \int \phi(v)\,dv+f(u) = f(u)+g(v)$.

15.02. We have

$$2\log\{\phi(x)+\psi(y)\}+z = \log 2+\log\phi'(x)+\log\psi'(y),$$

whence, keeping y constant and differentiating with respect to x,

$$2\phi'(x)/\{\phi(x)+\psi(y)\}+\frac{\partial z}{\partial x} = \phi''(x)/\phi'(x).$$

Differentiating this equation with respect to y, we find

$$-2\phi'(x)\psi'(y)/\{\phi(x)+\psi(y)\}^2+\frac{\partial^2 z}{\partial x\partial y}=0,\quad \text{i.e.}\quad \frac{\partial^2 z}{\partial x\partial y}=e^z.$$

15.03. $$\frac{1}{H}\frac{\partial H}{\partial x}=2\beta\frac{x}{t},\qquad -\frac{1}{H^2}\left(\frac{\partial H}{\partial x}\right)^2+\frac{1}{H}\frac{\partial^2 H}{\partial x^2}=\frac{2\beta}{t},$$

whence $$\frac{\partial^2 H}{\partial x^2}=2\beta\frac{H}{t}+4\beta^2\frac{x^2}{t^2}H.$$

Furthermore

$$\frac{1}{H}\frac{\partial H}{\partial t}=\alpha-\beta\frac{x^2}{t^2}-\frac{1}{2t}\quad\text{and so}\quad \frac{\partial^2 H}{\partial x^2}+4\beta\frac{\partial H}{\partial t}=4\alpha\beta H.$$

15.04. $$\frac{\partial H}{\partial r}=\frac{\partial H}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial H}{\partial y}\frac{\partial y}{\partial r}+\frac{\partial H}{\partial z}\frac{\partial z}{\partial r}$$

$$=\cos\theta\sin\phi\frac{\partial H}{\partial x}+\sin\theta\sin\phi\frac{\partial H}{\partial y}+\cos\phi\frac{\partial H}{\partial z},$$

$$\frac{\partial H}{\partial \phi}=\frac{\partial H}{\partial x}\frac{\partial x}{\partial \phi}+\frac{\partial H}{\partial y}\frac{\partial y}{\partial \phi}+\frac{\partial H}{\partial z}\frac{\partial z}{\partial \phi}$$

$$=r\cos\theta\cos\phi\frac{\partial H}{\partial x}+r\sin\theta\cos\phi\frac{\partial H}{\partial y}-r\sin\phi\frac{\partial H}{\partial z},$$

whence $$\sin\phi\frac{\partial H}{\partial r}+\frac{\cos\phi}{r}\frac{\partial H}{\partial \phi}=\cos\theta\frac{\partial H}{\partial x}+\sin\theta\frac{\partial H}{\partial y}.$$

Moreover

$$\frac{\partial H}{\partial \theta}=\frac{\partial H}{\partial x}\frac{\partial x}{\partial \theta}+\frac{\partial H}{\partial y}\frac{\partial y}{\partial \theta}+\frac{\partial H}{\partial z}\frac{\partial z}{\partial \theta}=-r\sin\theta\sin\phi\frac{\partial H}{\partial x}+r\cos\theta\sin\phi\frac{\partial H}{\partial y},$$

whence follows the second result.

15.05. Let $\phi(t)=f\{a+(A-a)t,b+(B-b)t\}$, so that $\phi(t)$ is differentiable in (0, 1), and $\phi(0)=\phi(1)$; hence by Rolle's theorem we can find τ in [0, 1] such that $\phi'(\tau)=0$. But

$$\phi'(t)=(A-a)f_x\{a+(A-a)t,b+(B-b)t\}+(B-b)f_y\{a+(A-a)t,b+(B-b)t\}$$

and so, writing $a+(A-a)\tau=\alpha$, $b+(B-b)\tau=\beta$, we have

$$(\alpha-a)f_x(\alpha,\beta)+(\beta-b)f_y(\alpha,\beta)=0,\quad \text{since } \tau\neq 0.$$

15.1. If $z=\frac{1}{2}\log(x^2+y^2)$ then

$$\frac{\partial z}{\partial x}=\frac{x}{x^2+y^2},\qquad \frac{\partial z}{\partial y}=\frac{y}{x^2+y^2}\quad\text{and so}\quad \left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2=1/(x^2+y^2);$$

if $z=\tan^{-1}(y/x)$ then

$$\frac{\partial z}{\partial x}=-\frac{y/x^2}{1+(y/x)^2}=-\frac{y}{x^2+y^2}\quad\text{and}\quad \frac{\partial z}{\partial y}=\frac{1/x}{1+(y/x)^2}=\frac{x}{x^2+y^2},\quad\text{etc.}$$

15.11. Regarding z as a function of x and y satisfying $\phi(x,y,z)=0$, we find $\frac{\partial\phi}{\partial x}+\frac{\partial\phi}{\partial z}\left(\frac{\partial z}{\partial x}\right)_y=0$; similarly, regarding y as a function of x and z,

$\dfrac{\partial\phi}{\partial z}+\dfrac{\partial\phi}{\partial y}\left(\dfrac{\partial y}{\partial z}\right)_x = 0$ and regarding x as a function of y and z, $\dfrac{\partial\phi}{\partial y}+\dfrac{\partial\phi}{\partial x}\left(\dfrac{\partial x}{\partial y}\right)_z = 0$; hence

$$\frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y}\frac{\partial\phi}{\partial z} = -\frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y}\frac{\partial\phi}{\partial z}\left\{\left(\frac{\partial z}{\partial x}\right)_y\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial z}\right)_x\right\}$$

and therefore
$$\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y\left(\frac{\partial x}{\partial y}\right)_z = -1,$$

since none of $\dfrac{\partial\phi}{\partial x}, \dfrac{\partial\phi}{\partial y}, \dfrac{\partial\phi}{\partial z}$ is zero when ϕ depends upon all three x, y, z.

15.2 Under the transformation $y = a_0+a_1x+a_2x^2+\ldots+a_{n-1}x^{n-1}+wx^n$, $P(x,y)$ takes the form

$$A_0+A_1x+A_2x^2+\ldots+A_{n-1}x^{n-1}+x^nQ(x,w),$$

where $Q(x,w)$ is a polynomial, and $A_0, A_1,\ldots, A_{n-1}$ are polynomials in the n variables $a_0, a_1,\ldots, a_{n-1}$. Thus a set of values $a_0, a_1,\ldots, a_{n-1}$ may be chosen so that $A_0, A_1,\ldots, A_{n-1}$ are all zero; for this set of values

$$P(x,y) = x^nQ(x,w),$$

whence, differentiating with respect to w, $\dfrac{\partial P}{\partial y}\dfrac{\partial y}{\partial w} = x^n\dfrac{\partial Q}{\partial w}$. Since $\dfrac{\partial y}{\partial w} = x^n$, it follows that for all values of x, except perhaps $x = 0$, $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial w}$; by continuity this holds also for $x = 0$. Let $w = w_0$ be a solution of $Q(0,w) = 0$, then since $\dfrac{\partial Q}{\partial w} = \dfrac{\partial P}{\partial y} \neq 0$ when $x = 0$, there is, by 15.85, a unique differentiable solution of the equation $Q(x,w) = 0$, say $w = w(x)$. If $a_n = w(0)$ and $\alpha = w(x)-w(0)$, so that $\alpha \to 0$ as $x \to 0$, then

$$y = a_0+a_1x+\ldots+a_{n-1}x^{n-1}+x^n(a_n+\alpha)$$

is a solution of $P(x,y) = 0$, where α is a differentiable function of x, such that $\alpha \to 0$, as $x \to 0$.

The solution may be written in the form $y = a_0+a_1x+\ldots+x^{n-1}(a_{n-1}+\beta)$, where $\beta = (a_n+\alpha)x$ so that $\beta \to 0$ as $x \to 0$ and we determine $a_0, a_1,\ldots, a_{n-1}$ by the condition that the coefficients of $x^0, x^1,\ldots, x^{n-1}$ in

$$P(x, a_0+a_1x+\ldots+a_{n-1}x^{n-1})$$

are all zero.

If $P(x,y) = 2x^2+xy-y^2+x^3-xy^2+y^3$, then $P(0,0) = 0$ and $P_y(0,0) = 0$ so there is not a unique solution at the origin; taking $y = zx$ we have

$$x^2P^*(x,z) = P(x,xz) = x^2\{2+z-z^2+x-xz^2+xz^3\}$$

and so
$$P_2^*(x,z) = 1-2z-2xz+3xz^2.$$

Now $P^*(0,z) = 0$ if $2+z-z^2 = 0$, i.e. $z = 2$ or $z = -1$, and $P_2^*(0,-1) = 3$, $P_2^*(0,2) = -3$ so that $P^*(x,z) = 0$ has a unique solution in the neighbourhood of $(0,2)$ and a unique solution in the neighbourhood of $(0,-1)$.

Each solution is of the form $z = a_0+a_1x+x^2(a_2+\beta_x)$, where $\beta_x \to 0$ as $x \to 0$ and a_0, a_1, a_2 are to be chosen so that the coefficients of x^0, x^1, x^2 in $P(x, a_0+a_1x+a_2x^2)$ are zero. Thus

$$2+a_0-a_0^2 = 0, \quad a_1-2a_0a_1+1-a_0^2+a_0^3 = 0, \quad a_2-a_1^2-2a_0a_2+3a_0^2a_1 = 0.$$

whence, either $a_0 = 2$, $a_1 = \frac{4}{3}$, $a_2 = -\frac{22}{9}$ or $a_0 = -1$, $a_1 = \frac{1}{3}$, $a_2 = -\frac{14}{9}$, giving the solutions $y = 2x+\frac{4}{3}x^2-x^3(\frac{22}{9}+\beta_x)$ and $y = -x+\frac{1}{3}x^2-x^3(\frac{14}{9}+\gamma_x)$, where β_x and γ_x tend to zero with x.

Furthermore $P(0,1) = 0$, $P_y(0,1) = 1$ and so there is a unique solution of $P(x,y)$ in the neighbourhood $(0,1)$ given by

$$y = c_0+c_1 x+c_2 x^2+x^3(c_3+\delta_x);$$

the coefficients c_0, c_1, c_2, c_3 are determined so that the coefficients of x^0, x^1, x^2, x^3 in $P(x, c_0+c_1 x+c_2 x^2+c_3 x^3)$ are all zero. Thus $-c_0^2+c_0^3 = 0$, $c_0-2c_0 c_1-c_0^2+3c_0^2 c_1 = 0$, $2+c_1-c_1^2-2c_0 c_2-2c_0 c_1+3c_0^2 c_2+3c_0 c_1^2 = 0$ and $c_2-2c_1 c_2+1-c_1^2-2c_0 c_3+3c_0^2 c_3+6c_0 c_1 c_2+c_1^3 = 0$; since $y = 1$ when $x = 0$, therefore $c_0 = 1$, whence $c_1 = 0$, $c_2 = -2$, $c_3 = -1$ and the required solution is

$$y = 1-2x^2-x^3(1+\delta_x).$$

15.21. Since $f(a,c)$ is the double limit of $f(x,y)$ as x and y tend to a and c, we can find α_k, γ_k such that

$$|f(x,y)-f(a,c)| < \frac{1}{2k} \quad \text{provided } a \leqslant x \leqslant \alpha_k,\ c \leqslant y \leqslant \gamma_k,$$

and therefore for any points (x,y), (X,Y) in $(a,\alpha_k)(c,\gamma_k)$

$$|f(X,Y)-f(x,y)| = |f(X,Y)-f(a,c)-\{f(x,y)-f(a,c)\}| < \frac{1}{k}.$$

In the interval $\alpha_k \leqslant x \leqslant b$, $f(x,c)$ is the continuous interval limit of $f(x,y)$ as $y \to c$, therefore $|f(x,y)-f(x,c)| < \frac{1}{3k}$ provided $c \leqslant y \leqslant y_k$, whence

$$|f(X,Y)-f(x,y)| = |f(X,Y)-f(X,c)+f(x,c)-f(x,y)+f(X,c)-f(x,c)|$$
$$< \frac{1}{k} \quad \text{provided } |X-x| \leqslant \lambda_k,\ c \leqslant y \leqslant Y \leqslant y_k,$$

since $|f(X,c)-f(x,c)| < \frac{1}{3k}$ provided $|X-x| \leqslant \lambda_k$, $f(x,c)$ being continuous. Thus the rectangle $(a,b)(c,y_k)$ may be divided into a finite number of parts such that for any two points (X,Y), (x,y) in the same part

$$|f(X,Y)-f(x,y)| < \frac{1}{k}.$$

Similarly we can find $x_k > a$, such that the rectangle $(a,x_k)(c,d)$ may be divided into a finite number of parts and

$$|f(X,Y)-f(x,y)| < \frac{1}{k}$$

for any (x,y), (X,Y) in the same part.

Furthermore, since $f(x,y)$ is continuous in the rectangle $(x_k,b)(y_k,d)$, this rectangle may be subdivided into a finite number of parts in each of which $|f(X,Y)-f(x,y)| < \frac{1}{k}$. Accordingly the whole rectangle $(a,b)(c,d)$ may be so divided, and therefore $f(x,y)$ is continuous in the whole rectangle.

15.3. $\dfrac{\partial z}{\partial x} = ae^{ax}f(x+y) - ae^{-ax}g(x-y) + e^{ax}f'(x+y) + e^{-ax}g'(x-y)$,

$$\frac{\partial^2 z}{\partial x^2} = a^2e^{ax}f(x+y) + a^2e^{-ax}g(x-y) + 2ae^{ax}f'(x+y)$$
$$-2ae^{-ax}g'(x-y) + e^{ax}f''(x+y) + e^{-ax}g''(x-y),$$

$\dfrac{\partial z}{\partial y} = e^{ax}f'(x+y) - e^{-ax}g'(x-y)$, and $\dfrac{\partial^2 z}{\partial y^2} = e^{ax}f''(x+y) + e^{-ax}g''(x-y)$,

whence the result follows.

15.31. $\dfrac{\partial r}{\partial x} = \dfrac{1}{2r}\dfrac{\partial r^2}{\partial x} = \dfrac{x}{r}$, $\dfrac{\partial r}{\partial y} = \dfrac{y}{r}$, whence $\dfrac{\partial^2 r}{\partial x^2} = \dfrac{1}{r} - \dfrac{x}{r^2}\dfrac{\partial r}{\partial x} = \dfrac{1}{r} - \dfrac{x^2}{r^3}$ and

$\dfrac{\partial^2 r}{\partial y^2} = \dfrac{1}{r} - \dfrac{y}{r^2}\dfrac{\partial r}{\partial y} = \dfrac{1}{r} - \dfrac{y^2}{r^3}$, and so $\nabla^2 r = \dfrac{2}{r} - \dfrac{x^2+y^2}{r^3} = \dfrac{1}{r}$.

Furthermore

$$\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = -\frac{1}{r^2}\frac{\partial r}{\partial x} = -\frac{x}{r^3}, \qquad \frac{\partial}{\partial y}\left(\frac{1}{r}\right) = -\frac{1}{r^2}\frac{\partial r}{\partial y} = -\frac{y}{r^3}$$

and therefore $\qquad \nabla^2\left(\dfrac{1}{r}\right) = -\dfrac{1}{r^3} + \dfrac{3x^2}{r^5} - \dfrac{1}{r^3} + \dfrac{3y^2}{r^5} = \dfrac{1}{r^3}$.

Since $\dfrac{\partial}{\partial x}\log r = \dfrac{1}{r}\dfrac{\partial r}{\partial x} = \dfrac{x}{r^2}$ and $\dfrac{\partial}{\partial y}\log r = \dfrac{y}{r^2}$ therefore

$$\nabla^2 \log r = \frac{2}{r^2} - \frac{2(x^2+y^2)}{r^4} = 0,$$

and since

$$\frac{\partial}{\partial x}\tan^{-1}\frac{y}{x} = -\frac{y}{x^2+y^2} = -\frac{y}{r^2}, \quad \frac{\partial}{\partial y}\tan^{-1}\frac{y}{x} = \frac{x}{x^2+y^2} = \frac{x}{r^2}$$

therefore $\qquad \nabla^2 \tan^{-1}\dfrac{y}{x} = \dfrac{2y}{r^3}\dfrac{x}{r} - \dfrac{2x}{r^3}\dfrac{y}{r} = 0.$

Writing $R^2 = x^2+y^2+z^2$, then $\dfrac{\partial R}{\partial x} = \dfrac{1}{2R}\dfrac{\partial R^2}{\partial x} = \dfrac{x}{R}$ and so $\dfrac{\partial^2 R}{\partial x^2} = \dfrac{1}{R} - \dfrac{x^2}{R^3}$,

whence $\qquad \left(\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}\right)R = \dfrac{3}{R} - \dfrac{x^2+y^2+z^2}{R^3} = \dfrac{2}{R}$.

Finally,

$\dfrac{\partial}{\partial x}\left(\dfrac{1}{R}\right) = -\dfrac{1}{R^2}\dfrac{\partial R}{\partial x} = -\dfrac{x}{R^3}$ and so $\dfrac{\partial^2}{\partial x^2}\left(\dfrac{1}{R}\right) = -\dfrac{1}{R^3} + \dfrac{3x}{R^4}\cdot\dfrac{x}{R} = -\dfrac{1}{R^3} + \dfrac{3x^2}{R^5}$

so that $\qquad \left(\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}\right)\dfrac{1}{R} = -\dfrac{3}{R^3} + \dfrac{3(x^2+y^2+z^2)}{R^5} = 0.$

15.4. A curve of the family $\Phi(x, y) = \lambda$ can be found through any given point, e.g. the curve $\Phi(x, y) = \Phi(a, b)$ passes through the point (a, b).

At a common point of a curve $\Phi(x, y) = \lambda$ and a curve $\Psi(x, y) = \mu$, the curves touch if dy/dx has the same value on each curve, i.e. if $\Phi_x + \Phi_y \dfrac{dy}{dx} = 0$

and $\Psi_x+\Psi_y\dfrac{dy}{dx}=0$ are true simultaneously, whence the locus of points of contact is

$$\Phi_x\Psi_y=\Phi_y\Psi_x.$$

The family of parabolas is $x^2/y=$ constant, and the family of circles is $\dfrac{(x-a)^2+y^2-b^2}{y}=$ constant, taking the x-coordinates of the given points to be $a-b$ and $a+b$.

The locus of points of contact is therefore

$$\frac{2x}{y}\left\{-\frac{(x-a)^2+y^2-b^2}{y^2}+2\right\}=-\frac{x^2}{y^2}\left\{\frac{2(x-a)}{y}\right\},$$

i.e. $y^2-(x-a)^2+b^2+x(x-a)=0$ or $y^2+a(x-a)+b^2=0$, which is a parabola.

15.41. On $\phi(x,y)=0$ we have

$$\phi_x+\phi_y\frac{dy}{dx}=0,\qquad \phi_{xx}+2\phi_{xy}\frac{dy}{dx}+\phi_{yy}\left(\frac{dy}{dx}\right)^2+\phi_y\frac{d^2y}{dx^2}=0,$$

whence
$$\phi_y^3\frac{d^2y}{dx^2}=-\{\phi_{xx}\phi_y^2-2\phi_{xy}\phi_x\phi_y+\phi_{yy}\phi_x^2\},$$

and so the centre of curvature at (x,y) is

$$\xi=x-\left\{1+\left(\frac{dy}{dx}\right)^2\right\}\frac{dy}{dx}\bigg/\frac{d^2y}{dx^2}=x-\phi_x(\phi_x^2+\phi_y^2)/(\phi_{xx}\phi_y^2-2\phi_{xy}\phi_x\phi_y+\phi_{yy}\phi_x^2),$$

$$\eta=y+\left\{1+\left(\frac{dy}{dx}\right)^2\right\}\bigg/\frac{d^2y}{dx^2}=y-\phi_y(\phi_x^2+\phi_y^2)/(\phi_{xx}\phi_y^2-2\phi_{xy}\phi_x\phi_y+\phi_{yy}\phi_x^2).$$

15.42. By 15.41

$$\frac{1}{\kappa^2}=(x-\xi)^2+(y-\eta)^2=(\phi_x^2+\phi_y^2)^3/(\phi_{xx}\phi_y^2-2\phi_{xy}\phi_x\phi_y+\phi_{yy}\phi_x^2)^2.$$

If $\phi(x,y)=\sin x+\sin y$ then $\phi_x=\cos x$, $\phi_{xx}=-\sin x$, $\phi_{xy}=0$, $\phi_y=\cos y$, and $\phi_{yy}=-\sin y$ and so

$$\kappa^2=\{\cos^2x+\cos^2y\}^{-3}\{\sin x\cos^2y+\sin y\cos^2x\}^2;$$

but, on the curve, $\sin x+\sin y=c$ and so

$$\cos^2x+\cos^2y=\cos^2x+1-(c-\sin x)^2=2\cos^2x+2c\sin x-c^2;$$

similarly
$$\sin x\cos^2y+\sin y\cos^2x=c(1+\sin^2x-c\sin x)$$

and therefore κ is the positive value of

$$c(1+\sin^2x-c\sin x)/(2\cos^2x+2c\sin x-c^2)^{\frac{3}{2}}.$$

15.43. Let the origin be the fixed point, and let $x\cos\alpha+y\sin\alpha=p$ be the directrix and (λ,μ) the focus of the variable parabola. Then the equation of the parabola is

$$2\phi(x,y)=(x-\lambda)^2+(y-\mu)^2-(x\cos\alpha+y\sin\alpha-p)^2=0,$$

which passes through the origin if $p^2=\lambda^2+\mu^2$. At the origin

$$\phi_x=p\cos\alpha-\lambda,\qquad \phi_y=p\sin\alpha-\mu,$$

$$\phi_{xx}=\sin^2\alpha,\qquad \phi_{xy}=-\sin\alpha\cos\alpha,\qquad \phi_{yy}=\cos^2\alpha.$$

If the inclination of the tangent at the origin is $-\delta$, where δ is constant, then

$$(p\cos\alpha-\lambda)/\cos\delta = (p\sin\alpha-\mu)/\sin\delta = v, \quad \text{say}.$$

Hence the constant radius of curvature ρ is given by

$$\rho = v^3/v^2\cos^2(\alpha-\delta) = v/\cos^2(\alpha-\delta).$$

Therefore

$$\lambda = p\cos\alpha-\rho\cos\delta\cos^2(\alpha-\delta), \qquad \mu = p\sin\alpha-\rho\sin\delta\cos^2(\alpha-\delta),$$

whence $\quad p^2 = \lambda^2+\mu^2 = p^2+\rho^2\cos^4(\alpha-\delta)-2p\rho\cos^3(\alpha-\delta)$

and so $\quad \rho\cos(\alpha-\delta) = 2p,$

i.e. $\quad \frac{1}{2}\rho\cos\delta\cos\alpha+\frac{1}{2}\rho\sin\delta\sin\alpha = p,$

which proves that the directrix $x\cos\alpha+y\sin\alpha = p$ passes through the fixed point $(\frac{1}{2}\rho\cos\delta, \frac{1}{2}\rho\sin\delta)$.

Furthermore, from

$$\frac{p\cos\alpha-\lambda}{\cos\delta} = \frac{p\sin\alpha-\mu}{\sin\delta} = \rho\cos^2(\alpha-\delta)$$

it follows that

$$p\cos(\alpha-\delta)-(\lambda\cos\delta+\mu\sin\delta) = \rho\cos^2(\alpha-\delta),$$

whence $\quad 2p^2-\rho(\lambda\cos\delta+\mu\sin\delta) = 4p^2,$

i.e. $\quad \lambda^2+\mu^2+\frac{1}{2}\rho(\lambda\cos\delta+\mu\sin\delta) = 0,$

which shows that (λ,μ) lies on a fixed circle.

15.44. Let $\phi(X,Y) \equiv (X-a)^2+(Y-b)^2 = c^2$ be the circle through P, P_1, P_2, where a, b, c are functions of x, x_1, x_2 the x-coordinates of P, P_1, P_2. The points $\{x,f(x)\}$, $\{x_1,f(x_1)\}$, $\{x_2,f(x_2)\}$ lie on the circle and so

$$\phi\{x,f(x)\} = \phi\{x_1,f(x_1)\} = \phi\{x_2,f(x_2)\}.$$

Hence by Rolle's theorem we can find α_1 in (x,x_1) and α_2 in (x,x_2) such that

$$\phi_x\{\alpha_1,f(\alpha_1)\}+\phi_y\{\alpha_1,f(\alpha_1)\}f'(\alpha_1) = 0,$$
$$\phi_x\{\alpha_2,f(\alpha_2)\}+\phi_y\{\alpha_2,f(\alpha_2)\}f'(\alpha_2) = 0$$

and hence, by another application of Rolle's theorem, this time to the function

$$\phi_x\{x,f(x)\}+\phi_y\{x,f(x)\}f'(x),$$

we can find γ in (α_1,α_2) such that

$$1+f'(\gamma)^2+(y-b)f''(\gamma) = 0 \quad (\text{since } \phi_{xx} = 2,\ \phi_{xy} = 0,\ \phi_{yy} = 2).$$

Hence as $x_1 \to x$, $x_2 \to x$, so that $\alpha_1 \to x$, $\alpha_2 \to x$, then a, b, c tend to values a_0, b_0, c_0 satisfying the equations

$$(x-a_0)^2+(y-b_0)^2 = c_0^2,$$
$$(x-a_0)+(y-b_0)f'(x) = 0,$$
$$1+f'(x)^2+(y-b_0)f''(x) = 0,$$

which are the conditions which determine the circle of curvature at (x,y) on $y = f(x)$.

15.5. Write $u+v = w$, and treat x, u, w as the independent variables and v and y as functions of x, u, w; then since $f(y,v)-f(x,w) = 0$ we have, differentiating with respect to u,

$$\frac{\partial f(y,v)}{\partial y}\frac{\partial y}{\partial u}+\frac{\partial f(y,v)}{\partial v}\frac{\partial v}{\partial u} = 0,$$

and differentiating $u+v = w$ with respect to u, it follows that $1+\dfrac{\partial v}{\partial u} = 0$, whence

$$\frac{\partial y}{\partial u} = \frac{\partial f(y,v)}{\partial v}\bigg/\frac{\partial f(y,v)}{\partial y}.$$

But $y = f(x,u)$, so that $\dfrac{\partial y}{\partial u} = \dfrac{\partial f(x,u)}{\partial u}$, and therefore $\dfrac{\partial f(y,v)}{\partial v}\bigg/\dfrac{\partial f(y,v)}{\partial y}$ is independent of v, so that $\partial y/\partial u$ is a function of y alone.

15.51.

$$\begin{vmatrix} \alpha & \beta & f \\ \alpha_x & \beta_x & f_x \\ \alpha_y & \beta_y & f_y \end{vmatrix} = \begin{vmatrix} \alpha & \beta & f-\alpha f_\alpha-\beta f_\beta \\ \alpha_x & \beta_x & f_x-\alpha_x f_\alpha-\beta_x f_\beta \\ \alpha_y & \beta_y & f_y-\alpha_y f_\alpha-\beta_y f_\beta \end{vmatrix} = (f-\alpha f_\alpha-\beta f_\beta)\frac{\partial(\alpha,\beta)}{\partial(x,y)},$$

since $f_x = f_\alpha \alpha_x+f_\beta \beta_x$ and $f_y = f_\alpha \alpha_y+f_\beta \beta_y$.

15.52. Since $S \equiv \alpha\beta-\frac{1}{2}\gamma^2 = 0$, therefore $S_\alpha \alpha_t+S_\beta \beta_t = 0$, regarding γ as a function of α, β, and so

$$\frac{\alpha_t}{S_\beta} = \frac{\beta_t}{-S_\alpha} = \frac{\beta\alpha_t-\alpha\beta_t}{\alpha S_\alpha+\beta S_\beta} = \frac{\beta\alpha_t-\alpha\beta_t}{\gamma(\gamma-\alpha\gamma_\alpha-\beta\gamma_\beta)}.$$

But $x_t = x_\alpha \alpha_t+x_\beta \beta_t$, regarding x as a function of α, β, and therefore

$$\frac{\alpha_t}{S_\beta} = \frac{\beta_t}{-S_\alpha} = x_t\bigg/\begin{vmatrix} x_\alpha & S_\alpha \\ x_\beta & S_\beta \end{vmatrix} = x_t\bigg/\frac{\partial(x,S)}{\partial(\alpha,\beta)} = \left(x_t\bigg/\frac{\partial(x,S)}{\partial(x,y)}\right)\frac{\partial(\alpha,\beta)}{\partial(x,y)} = \frac{x_t}{S_y}\frac{\partial(\alpha,\beta)}{\partial(x,y)},$$

which completes the proof.

15.521.

$$\frac{\partial}{\partial x}\begin{vmatrix} \alpha & \beta & \gamma \\ \alpha_x & \beta_x & \gamma_x \\ \alpha_y & \beta_y & \gamma_y \end{vmatrix} = \begin{vmatrix} \alpha_x & \beta_x & \gamma_x \\ \alpha_x & \beta_x & \gamma_x \\ \alpha_y & \beta_y & \gamma_y \end{vmatrix} + \begin{vmatrix} \alpha & \beta & \gamma \\ 0 & 0 & 0 \\ \alpha_y & \beta_y & \gamma_y \end{vmatrix} + \begin{vmatrix} \alpha & \beta & \gamma \\ \alpha_x & \beta_x & \gamma_x \\ 0 & 0 & 0 \end{vmatrix} = 0,$$

and similarly

$$\frac{\partial}{\partial y}\begin{vmatrix} \alpha & \beta & \gamma \\ \alpha_x & \beta_x & \gamma_x \\ \alpha_y & \beta_y & \gamma_y \end{vmatrix} = 0.$$

15.53. If $S = 0$ is the conic, $\gamma = 0$ the chord, and $\alpha = 0$, $\beta = 0$ the tangents at its ends then α, β can be chosen so that $S = \alpha\beta-\frac{1}{2}\gamma^2$.

By 15.52,

$$\frac{x_t}{\gamma S_y}\frac{\partial(\alpha,\beta)}{\partial(x,y)} = \frac{\beta\alpha_t-\alpha\beta_t}{\gamma^2(\gamma-\alpha\gamma_\alpha-\beta\gamma_\beta)} = \frac{1}{2(\gamma-\alpha\gamma_\alpha-\beta\gamma_\beta)}\left(\frac{\alpha_t}{\alpha}-\frac{\beta_t}{\beta}\right).$$

But, by 15.51,

$$(\gamma-\alpha\gamma_{\alpha}-\beta\gamma_{\beta})\frac{\partial(\alpha,\beta)}{\partial(x,y)}=\begin{vmatrix}\alpha & \beta & \gamma\\ \alpha_x & \beta_x & \gamma_x\\ \alpha_y & \beta_y & \gamma_y\end{vmatrix}=\text{constant},$$

since α, β, γ are linear functions of x, y. Hence, since $\gamma = lx+my+n$ and $S_y = 2(hx+by+f)$, therefore

$$\int\frac{x_t\,dt}{(lx+my+n)(hx+by+f)}=A\log\frac{\alpha}{\beta}+B,$$

where A and B are constants, whence the result follows, since the perpendiculars from a current point P to the tangents $\alpha=0$, $\beta=0$ are proportional to α and β.

15.6. We have $\xi = u+v+w = 2x+z$, $\eta = uv+vw+wu = 3x^2-yz$, $\zeta = uvw = z^4/y$, and

$$\frac{\partial(u,v,w)}{\partial(x,y,z)}=\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}\Big/\frac{\partial(\xi,\eta,\zeta)}{\partial(u,v,w)}.$$

Thus $$\frac{\partial(\xi,\eta,\zeta)}{\partial(u,v,w)}=\begin{vmatrix}1 & v+w & vw\\ 1 & w+u & wu\\ 1 & v+u & uv\end{vmatrix}=-(u-v)(v-w)(w-u),$$

and $$\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}=\begin{vmatrix}2 & 6x & 0\\ 0 & -z & -z^4/y^2\\ 1 & -y & 4z^3/y\end{vmatrix}=-2(z^4/y^2)\begin{vmatrix}1 & 3x & 0\\ 0 & 1 & 1\\ 1 & -y & 4y\end{vmatrix}$$

$$=-2(3x+5y)z^4/y^2,$$

which completes the proof.

15.61. $$\frac{\partial H}{\partial x}=\frac{\partial H}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial H}{\partial v}\frac{\partial v}{\partial x}=2\left(x\frac{\partial H}{\partial u}+y\frac{\partial H}{\partial v}\right),$$

$$\frac{\partial H}{\partial y}=2\left(-y\frac{\partial H}{\partial u}+x\frac{\partial H}{\partial v}\right),$$

and so $$\frac{\partial^2 H}{\partial x^2}=2\frac{\partial H}{\partial u}+4x\left(x\frac{\partial}{\partial u}+y\frac{\partial}{\partial v}\right)\frac{\partial H}{\partial u}+4y\left(x\frac{\partial}{\partial u}+y\frac{\partial}{\partial v}\right)\frac{\partial H}{\partial v}$$

$$=2\frac{\partial H}{\partial u}+4\left(x^2\frac{\partial^2}{\partial u^2}+2xy\frac{\partial^2}{\partial u\partial v}+y^2\frac{\partial^2}{\partial v^2}\right)H,$$

and similarly

$$\frac{\partial^2 H}{\partial y^2}=-2\frac{\partial H}{\partial u}+4\left(y^2\frac{\partial^2}{\partial u^2}-2xy\frac{\partial^2}{\partial u\partial v}+x^2\frac{\partial^2}{\partial v^2}\right)H,$$

whence

$$y^2\frac{\partial^2 H}{\partial x^2}-x^2\frac{\partial^2 H}{\partial y^2}=2(y^2+x^2)\frac{\partial H}{\partial u}+8xy(y^2+x^2)\frac{\partial^2 H}{\partial u\partial v}+4(y^4-x^4)\frac{\partial^2 H}{\partial v^2};$$

and $$x\frac{\partial H}{\partial x}-y\frac{\partial H}{\partial y} = 2(y^2+x^2)\frac{\partial H}{\partial u},$$
so that the transformed equation is
$$2xy\frac{\partial^2 H}{\partial u\partial v}+(y^2-x^2)\frac{\partial^2 H}{\partial v^2} = 0, \quad \text{i.e.} \quad \left(u\frac{\partial}{\partial v}-v\frac{\partial}{\partial u}\right)\frac{\partial H}{\partial v} = 0.$$

15.62. We have
$$x = \tfrac{1}{2}\{(u+v)^n+(u-v)^n\}, \qquad y = \tfrac{1}{2}\{(u+v)^n-(u-v)^n\},$$
and
$$\frac{\partial^2 f}{\partial u^2} = \frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial u}\right)^2+2\frac{\partial^2 f}{\partial x\partial y}\left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial y}{\partial u}\right)+\frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial u}\right)^2+\frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial u^2}+\frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial u^2},$$
$$\frac{\partial^2 f}{\partial v^2} = \frac{\partial^2 f}{\partial x^2}\left(\frac{\partial x}{\partial v}\right)^2+2\frac{\partial^2 f}{\partial x\partial y}\left(\frac{\partial x}{\partial v}\right)\left(\frac{\partial y}{\partial v}\right)+\frac{\partial^2 f}{\partial y^2}\left(\frac{\partial y}{\partial v}\right)^2+\frac{\partial f}{\partial x}\frac{\partial^2 x}{\partial v^2}+\frac{\partial f}{\partial y}\frac{\partial^2 y}{\partial v^2}.$$
But $$\frac{\partial x}{\partial u} = \frac{n}{2}\{(u+v)^{n-1}+(u-v)^{n-1}\} = \frac{\partial y}{\partial v}$$
and $$\frac{\partial x}{\partial v} = \frac{n}{2}\{(u+v)^{n-1}-(u-v)^{n-1}\} = \frac{\partial y}{\partial u},$$
so that $$\frac{\partial^2 x}{\partial u^2} = \frac{\partial^2 y}{\partial u\partial v} = \frac{\partial^2 x}{\partial v^2} \quad \text{and} \quad \frac{\partial^2 y}{\partial u^2} = \frac{\partial^2 y}{\partial v^2},$$
and therefore
$$\frac{\partial^2 f}{\partial u^2}-\frac{\partial^2 f}{\partial v^2} = \left(\frac{\partial^2 f}{\partial x^2}-\frac{\partial^2 f}{\partial y^2}\right)\left\{\left(\frac{\partial x}{\partial u}\right)^2-\left(\frac{\partial x}{\partial v}\right)^2\right\}$$
$$= \left(\frac{\partial^2 f}{\partial x^2}-\frac{\partial^2 f}{\partial y^2}\right)n^2(u+v)^{n-1}(u-v)^{n-1}$$
$$= \left(\frac{\partial^2 f}{\partial x^2}-\frac{\partial^2 f}{\partial y^2}\right)n^2\frac{(u^2-v^2)^n}{u^2-v^2},$$
i.e. $$(u^2-v^2)\left(\frac{\partial^2 f}{\partial u^2}-\frac{\partial^2 f}{\partial v^2}\right) = n^2(x^2-y^2)\left(\frac{\partial^2 f}{\partial x^2}-\frac{\partial^2 f}{\partial y^2}\right).$$

15.621. Regarding y and z as functions of x and w we have, from $z = z(x,y)$ and $w = w(x,y)$,
$$\left(\frac{\partial z}{\partial x}\right)_w = \left(\frac{\partial z}{\partial x}\right)_y+\left(\frac{\partial z}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_w \quad \text{and} \quad 0 = \left(\frac{\partial w}{\partial x}\right)_y+\left(\frac{\partial w}{\partial y}\right)_x\left(\frac{\partial y}{\partial x}\right)_w,$$
whence $$\left(\frac{\partial z}{\partial x}\right)_w\left(\frac{\partial w}{\partial y}\right)_x = \left(\frac{\partial z}{\partial x}\right)_y\left(\frac{\partial w}{\partial y}\right)_x-\left(\frac{\partial z}{\partial y}\right)_x\left(\frac{\partial w}{\partial x}\right)_y.$$

15.63. Regard x and y as functions of t satisfying $u^2 = 2kxy$ for all values of t; then, differentiating with respect to t,
$$2u(u_x x_t+u_y y_t) = 2k(xy_t+yx_t);$$
but $u_x = a$, $u_y = b$, and so
$$\frac{x_t}{ub-kx} = \frac{y_t}{ky-ua} = \frac{xy_t-yx_t}{kxy-uax+kxy-uby} = \frac{xy_t-yx_t}{u(u-ax-by)},$$

and so
$$\frac{x_1}{u(ub-kx)} = \frac{xy_1-yx_1}{2ckxy} = \frac{1}{2ck}\left(\frac{y_1}{y}-\frac{x_1}{x}\right),$$
which completes the proof. Example 15.63 is of course a special case of Example 15.53.

15.7. $$\frac{\partial H}{\partial u} = \frac{\partial H}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial H}{\partial y}\frac{\partial y}{\partial u} = \frac{\partial H}{\partial x}\cos\phi+\frac{\partial H}{\partial y}\sin\phi,$$
$$\frac{\partial^2 H}{\partial u^2} = \frac{\partial^2 H}{\partial x^2}\cos^2\phi+2\frac{\partial^2 H}{\partial x\partial y}\cos\phi\sin\phi+\frac{\partial^2 H}{\partial y^2}\sin^2\phi,$$
$$\frac{\partial^2 H}{\partial \phi^2} = \frac{\partial^2 H}{\partial x^2}u^2\sin^2\phi-2\frac{\partial^2 H}{\partial x\partial y}u^2\sin\phi\cos\phi+\frac{\partial^2 H}{\partial y^2}u^2\cos^2\phi-$$
$$-\frac{\partial H}{\partial x}u\cos\phi-\frac{\partial H}{\partial y}u\sin\phi,$$
whence
$$\frac{\partial^2 H}{\partial u^2}+\frac{1}{u^2}\frac{\partial^2 H}{\partial \phi^2}+\frac{1}{u}\frac{\partial H}{\partial u} = \frac{\partial^2 H}{\partial x^2}+\frac{\partial^2 H}{\partial y^2}.$$
Write $u = r\sin\theta$, then $x = u\cos\phi$ and $y = u\sin\phi$, and so
$$\frac{\partial^2 H}{\partial x^2}+\frac{\partial^2 H}{\partial y^2}+\frac{\partial^2 H}{\partial z^2} = \frac{\partial^2 H}{\partial u^2}+\frac{1}{u^2}\frac{\partial^2 H}{\partial \phi^2}+\frac{1}{u}\frac{\partial H}{\partial u}+\frac{\partial^2 H}{\partial z^2},$$
and similarly, since $u = r\sin\theta$, $z = r\cos\theta$, therefore
$$\frac{\partial^2 H}{\partial u^2}+\frac{\partial^2 H}{\partial z^2} = \frac{\partial^2 H}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 H}{\partial \theta^2}+\frac{1}{r}\frac{\partial H}{\partial r}$$
and
$$\frac{\partial H}{\partial u} = \frac{\partial H}{\partial r}\frac{\partial r}{\partial u}+\frac{\partial H}{\partial \theta}\frac{\partial \theta}{\partial u} = \frac{\partial H}{\partial r}\frac{u}{r}+\frac{\partial H}{\partial \theta}\frac{z}{r^2}, \quad \text{since } r^2 = u^2+z^2,\ \theta = \tan^{-1}u/z.$$
Thus
$$\frac{\partial^2 H}{\partial x^2}+\frac{\partial^2 H}{\partial y^2}+\frac{\partial^2 H}{\partial z^2} = \frac{\partial^2 H}{\partial r^2}+\frac{1}{r^2}\frac{\partial^2 H}{\partial \theta^2}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 H}{\partial \phi^2}+\frac{2}{r}\frac{\partial H}{\partial r}+\frac{\cot\theta}{r^2}\frac{\partial H}{\partial \theta}.$$

15.71. Since x, y may be regarded as functions of p and q therefore Z is a function of X and Y.

Hence
$$P = \frac{\partial Z}{\partial X} = \frac{\partial}{\partial p}(px+qy-z) = x+p\frac{\partial x}{\partial p}+q\frac{\partial y}{\partial p}-\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial p}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial p}\right) = x,$$
and similarly $Q = y$. Regarding X, Y as functions of P and Q, we have
$$\frac{\partial X}{\partial P} = \frac{\partial p}{\partial x} = r, \qquad \frac{\partial X}{\partial Q} = \frac{\partial^2 z}{\partial x\partial y} = s, \qquad \frac{\partial Y}{\partial P} = s, \qquad \frac{\partial Y}{\partial Q} = \frac{\partial q}{\partial y} = t.$$
Hence
$$rR+sS = \frac{\partial X}{\partial P}\cdot\frac{\partial P}{\partial X}+\frac{\partial X}{\partial Q}\cdot\frac{\partial Q}{\partial X} = 1$$
and
$$rS+sT = \frac{\partial P}{\partial Y}\cdot\frac{\partial X}{\partial P}+\frac{\partial Q}{\partial Y}\cdot\frac{\partial X}{\partial Q} = 0.$$
Similarly $sR+tS = 0$, $sS+tT = 1$, and therefore
$$\frac{r}{T} = -\frac{s}{S} = \frac{t}{R} = \frac{rt-s^2}{Tt+Ss} = rt-s^2;$$

15.72. We have

$$px+qy+rz = x\frac{\partial T}{\partial x}+y\frac{\partial T}{\partial y}+z\frac{\partial T}{\partial z} = 2T,$$

since T is homogeneous and of the second degree in x, y, z,

i.e. $$px+qy+rz = 2T.$$

Since $p = ax+hy+gz$, $q = hx+by+fz$, $r = gx+fy+cz$, and the determinant Δ is not identically zero, therefore we can solve for x, y, z in terms of p, q, and r (as well as u, v).

Regarding x, y, z as functions of the five independent variables p, q, r, u, and v we have

$$\left(\frac{\partial T}{\partial u}\right)_{p,q,r} = \left(\frac{\partial T}{\partial u}\right)_{x,y,z}+\frac{\partial T}{\partial x}\left(\frac{\partial x}{\partial u}\right)_{p,q,r}+\frac{\partial T}{\partial y}\left(\frac{\partial y}{\partial u}\right)_{p,q,r}+\frac{\partial T}{\partial z}\left(\frac{\partial z}{\partial u}\right)_{p,q,r}$$

$$= \left(\frac{\partial T}{\partial u}\right)_{x,y,z}+p\left(\frac{\partial x}{\partial u}\right)_{p,q,r}+q\left(\frac{\partial y}{\partial u}\right)_{p,q,r}+r\left(\frac{\partial z}{\partial u}\right)_{p,q,r}.$$

But, differentiating (i) with respect to u,

$$2\left(\frac{\partial T}{\partial u}\right)_{p,q,r} = p\left(\frac{\partial x}{\partial u}\right)_{p,q,r}+q\left(\frac{\partial y}{\partial u}\right)_{p,q,r}+r\left(\frac{\partial z}{\partial u}\right)_{p,q,r},$$

and therefore $$\left(\frac{\partial T}{\partial u}\right)_{p,q,r} = -\left(\frac{\partial T}{\partial u}\right)_{x,y,z}.$$

Similarly $$\left(\frac{\partial T}{\partial v}\right)_{p,q,r} = -\left(\frac{\partial T}{\partial v}\right)_{x,y,z}.$$

Furthermore

$$\left(\frac{\partial T}{\partial p}\right)_{q,r} = \left(\frac{\partial T}{\partial x}\right)_{y,z}\left(\frac{\partial x}{\partial p}\right)_{q,r}+\left(\frac{\partial T}{\partial y}\right)_{z,x}\left(\frac{\partial y}{\partial p}\right)_{q,r}+\left(\frac{\partial T}{\partial z}\right)_{x,y}\left(\frac{\partial z}{\partial p}\right)_{q,r},$$

and from (i) $$2\left(\frac{\partial T}{\partial p}\right)_{q,r} = x+p\left(\frac{\partial x}{\partial p}\right)_{q,r}+q\left(\frac{\partial y}{\partial p}\right)_{q,r}+r\left(\frac{\partial z}{\partial p}\right)_{q,r},$$

whence, since $p = \dfrac{\partial T}{\partial x}$, etc., we have $\left(\dfrac{\partial T}{\partial p}\right)_{q,r} = x$. Similarly $\left(\dfrac{\partial T}{\partial q}\right)_{p,r} = y$ and $\left(\dfrac{\partial T}{\partial r}\right)_{p,q} = z$.

15.8. Let x, y, z be the roots of the equation, then $\xi = x+y+z$, $\eta = xy+yz+zx$, $\zeta = xyz$, and so

$$\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)} = \begin{vmatrix} 1 & y+z & yz \\ 1 & z+x & zx \\ 1 & x+y & xy \end{vmatrix} = -(x-y)(y-z)(z-x)$$

and

$$\frac{\partial(s_{n+1}, s_{n+2}, s_{n+3})}{\partial(x,y,z)} = \begin{vmatrix} (n+1)x^n & (n+2)x^{n+1} & (n+3)x^{n+2} \\ (n+1)y^n & (n+2)y^{n+1} & (n+3)y^{n+2} \\ (n+1)z^n & (n+2)z^{n+1} & (n+3)z^{n+2} \end{vmatrix},$$

$$= (n+1)(n+2)(n+3)x^n y^n z^n \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$$

$$= (n+1)(n+2)(n+3)\zeta^n(x-y)(y-z)(z-x),$$

and therefore

$$\frac{\partial(s_{n+1}, s_{n+2}, s_{n+3})}{\partial(\xi, \eta, \zeta)} = \frac{\partial(s_{n+1}, s_{n+2}, s_{n+3})}{\partial(x, y, z)} \Big/ \frac{\partial(\xi, \eta, \zeta)}{\partial(x, y, z)}$$
$$= -(n+1)(n+2)(n+3)\zeta^n.$$

15.9. If $u = l(xy)$, $v = l(x)+l(y)$ then $u_x = yl'(xy) = y/xy = 1/x$; $x > 0$, $y > 0$. Similarly $u_y = 1/y$, $v_x = 1/x$, $v_y = 1/y$, whence $\dfrac{\partial(u,v)}{\partial(x,y)} = \dfrac{1}{xy} - \dfrac{1}{xy} = 0$ so that there is a functional relation connecting u and v, say $v = \phi(u)$, i.e. $l(x)+l(y) = \phi\{l(xy)\}$; hence, if $y = 1$, $l(x) = \phi\{l(x)\}$ for all $x > 0$, and therefore $l(x)+l(y) = \phi\{l(xy)\} = l(xy)$, provided $x > 0$, $y > 0$.

15.91. If $u = f(x+y)$, $v = f(x)f(y)$ then $u_x = f'(x+y) = u$, $u_y = f'(x+y) = u$ and $v_x = f'(x)f(y) = v$, $v_y = f(x)f'(y) = v$, whence $\dfrac{\partial(u,v)}{\partial(x,y)} = uv-uv = 0$ and so $v = \phi(u)$, i.e. $f(x)f(y) = \phi\{f(x+y)\}$ for all x and y. Taking $y = 0$ we have, for all x, $f(x) = \phi\{f(x)\}$, and therefore $f(x)f(y) = \phi\{f(x+y)\} = f(x+y)$.

15.92. If

$$u = f(x+y), \quad v = f(x)g(y)+f(y)g(x) \quad \text{then} \quad u_x = f'(x+y) = g(x+y),$$
$$u_y = g(x+y) \quad \text{and} \quad v_x = g(x)g(y)-f(x)f(y), \quad v_y = g(x)g(y)-f(x)f(y)$$

whence $\dfrac{\partial(u,v)}{\partial(x,y)} = 0$ for all x and y. Accordingly

$$f(x)g(y)+f(y)g(x) = \phi\{f(x+y)\}$$

for all x and y; taking $y = 0$ we have, for all x, $f(x) = \phi\{f(x)\}$, and therefore

$$f(x)g(y)+f(y)g(x) = \phi\{f(x+y)\} = f(x+y), \quad \text{etc.}$$

15.93. $$\int_a^{a^2} \frac{1}{x^2+a^2}\,dx = \left[\frac{1}{a}\tan^{-1}\frac{x}{a}\right]_a^{a^2} = \frac{1}{a}[\tan^{-1}a - \tfrac{1}{4}\pi];$$

hence $$\frac{d}{da}\int_a^{a^2} \frac{1}{x^2+a^2}\,dx = -\frac{1}{a^2}[\tan^{-1}a - \tfrac{1}{4}\pi] + \frac{1}{a(1+a^2)}.$$

But $$\frac{d}{da}\int_a^{a^2} \frac{1}{x^2+a^2}\,dx = -\int_a^{a^2} \frac{2a}{(x^2+a^2)^2}\,dx + \frac{2a}{a^4+a^2} - \frac{1}{a^2+a^2},$$

whence $$\int_a^{a^2} \frac{1}{(x^2+a^2)^2}\,dx = \frac{1}{2a^3}\tan^{-1}a - \frac{\pi}{8a^3} - \frac{(1-a)^2}{4a^2(1+a^2)}.$$

Furthermore

$$\frac{d}{da}\int_a^{a^2} \frac{1}{(x^2+a^2)^2}\,dx = -\frac{3}{2a^4}\tan^{-1}a + \frac{1}{2a^3(1+a^2)} + \frac{3\pi}{8a^4} - \frac{1}{a^3} + \frac{a}{(1+a^2)^2} + \frac{3}{4a^4}$$

and $$\frac{d}{da}\int_a^{a^2}\frac{1}{(x^2+a^2)^2}\,dx = -\int_a^{a^2}\frac{4a}{(x^2+a^2)^3}\,dx+\frac{2a}{(a^4+a^2)^2}-\frac{1}{4a^4},$$

whence

$$\int_a^{a^2}\frac{1}{(x^2+a^2)^2}\,dx = \frac{3}{8a^5}\tan^{-1}a-\frac{3\pi}{32a^5}-\frac{(a-1)(2a^3-a^2+3a-2)}{8a^4(1+a^2)^2}.$$

15.931. Since

$$\frac{d}{dx}\int_0^x (x-t)^{m-1}f(t)\,dt = \int_0^x \frac{\partial}{\partial x}(x-t)^{m-1}f(t)\,dt+\frac{dx}{dx}(x-x)^{m-1}f(x)$$

$$= (m-1)\int_0^x (x-t)^{m-2}f(t)\,dt$$

therefore $$\frac{d^{m-1}}{dx^{m-1}}\int_0^x (x-t)^{m-1}f(t)\,dt = (m-1)!\int_0^x f(t)\,dt,$$

whence $$\frac{d^m}{dx^m}\frac{1}{(m-1)!}\int_0^x (x-t)^{m-1}f(t)\,dt = \frac{d}{dx}\int_0^x f(t)\,dt = f(x).$$

15.932. Since $f(x, y)$ is continuous in R, we can find k so that $|f(x,y)| < k$ in R.

Let R be $(x_0, x_1)(y_0, y_1)$ and choose (x_0^*, x_1^*) containing a, and contained in (x_0, x_1) such that if $|y-b| < |x-a|k$, and x lies in (x_0^*, x_1^*), then y lies in (y_0, y_1). Call $(x_0^*, x_1^*)(y_0, y_1)$ the rectangle R^*.

We prove first that if x lies in (x_0^*, x_1^*) then $\phi_n(x)$ lies in (y_0, y_1), for any n. For $|\phi_1(x)-b| = \left|\int_a^x f(x,b)\,dx\right| < |x-a|k$, so that $\phi_1(x)$ lies in (y_0, y_1), and if $\phi_p(x)$ lies in (y_0, y_1) then $|\phi_{p+1}(x)-b| = \left|\int_a^x f\{x,\phi_p(x)\}\,dx\right| < |x-a|k$, which proves that $\phi_{p+1}(x)$ lies in (y_0, y_1) whence, by induction, $\phi_n(x)$ lies in (y_0, y_1) for all n.

Let μ be a bound of $\left|\int_a^x |f(t,b)|\,dt\right|$ in R^*; since

$$|f(x,Y)-f(x,y)| < M|Y-y|$$

it follows that

$$|\phi_{n+2}(x)-\phi_{n+1}(x)| = \left|\int_a^x [f\{x,\phi_{n+1}(x)\}-f\{x,\phi_n(x)\}]\,dx\right|$$

$$\leqslant M\left|\int_a^x |\phi_{n+1}(x)-\phi_n(x)|\,dx\right|.$$

Hence $\quad |\phi_2(x)-\phi_1(x)| \leqslant M\left|\int_a^x |\phi_1(x)-\phi_0(x)|\,dx\right| < M\mu|x-a|,$

and

$$|\phi_3(x)-\phi_2(x)| \leqslant M\left|\int_a^x |\phi_2(x)-\phi_1(x)|\,dx\right| \leqslant \mu M^2\left|\int_a^x |x-a|\,dx\right| = \mu M^2\frac{(x-a)^2}{2!},$$

for if $x > a$, $\quad \left|\int_a^x |x-a|\,dx\right| = \int_a^x (x-a)\,dx = \frac{(x-a)^2}{2};$

and if $x < a$, then

$$\left|\int_a^x |x-a|\,dx\right| = \int_x^a (a-x)\,dx = \frac{(a-x)^2}{2} = \frac{(x-a)^2}{2};$$

and if, for some p, $\quad |\phi_{p+1}(x)-\phi_p(x)| < \mu M^p \dfrac{|x-a|^p}{p!},$

then

$$|\phi_{p+2}(x)-\phi_{p+1}(x)| < \mu\frac{M^{p+1}}{p!}\left|\int_a^x |x-a|^p\,dx\right| = \mu\frac{M^{p+1}|x-a|^{p+1}}{(p+1)!}.$$

Thus $\quad |\phi_{n+1}(x)-\phi_n(x)| < \mu M^p \dfrac{|x-a|^p}{p!}$ for any n.

Since $\sum \dfrac{M^p|x-a|^p}{p!} = e^{M|x-a|}$, is interval-convergent in any interval, therefore $\sum |\phi_{n+1}(x)-\phi_n(x)|$ is interval-convergent in R^*, and therefore $\sum \{\phi_{n+1}(x)-\phi_n(x)\}$ is interval-convergent in R^*, i.e. the sequence $\phi_n(x)$ is interval-convergent in R^*.

But $\quad |f\{x,\phi_{n+1}(x)\}-f\{x,\phi_n(x)\}| < M|\phi_{n+1}(x)-\phi_n(x)|,$

and so $\sum |f\{x,\phi_{n+1}(x)\}-f\{x,\phi_n(x)\}|$ is interval-convergent, whence the sequence $f\{x,\phi_n(x)\}$ is interval-convergent.

Furthermore $\dfrac{d}{dx}\phi_{n+1}(x) = f\{x,\phi_n(x)\}$, and therefore, if $\psi(x) = \lim \phi_n(x)$, it follows by Theorem 14.61 that $\dfrac{d}{dx}\psi(x) = f\{x,\psi(x)\}$, since $f(x,y)$ is continuous, and so $y = \psi(x)$ is a solution of the equation $\dfrac{dy}{dx} = f(x,y)$; furthermore, $\phi_n(a) = b$, for all n, so that $\psi(a) = b$.

It remains to prove that $y = \psi(x)$ is the only solution in R^* of the differential equation which takes the value b at $x = a$.

Let $y = \omega(x)$ be a solution of the differential equation in R^* taking the value b at $x = a$. Divide the interval (x_0^*, x_1^*) into p equal parts by the points $x_0^* = \xi_0, \xi_1, \xi_2,\ldots, \xi_p = x_1^*$, where p is chosen so that the length of each part is less than $1/2M$.

We show first that if $\psi(x) = \omega(x)$ at some point of a closed interval (ξ_r, ξ_{r+1}) then $\psi(x) = \omega(x)$ throughout the interval.

Let $\psi(x) = \omega(x)$ at $x = \alpha$ in (ξ_r, ξ_{r+1}), then if $g(x) = \psi(x) - \omega(x)$,

$$|g(x)| = \left|\int_\alpha^x \{\psi'(t) - \omega'(t)\}\, dt\right| = \left|\int_\alpha^x [f\{t, \psi(t)\} - f\{t, \omega(t)\}]\, dt\right|$$

$$< M\left|\int_\alpha^x |\psi(t) - \omega(t)|\, dt\right| = M\left|\int_\alpha^x |g(t)|\, dt\right|$$

$$= M|x-\alpha|\,|g(c_1)|, \quad \text{for a } c_1 \text{ in } (\alpha, x), \text{ by the mean-value theorem,}$$

$$< \tfrac{1}{2}|g(c_1)| \quad \text{in } (\xi_r, \xi_{r+1}), \text{ since } |x-\alpha| < |\xi_{r+1} - \xi_r| < 1/2M.$$

Hence if c_2 is the value of $c_1(x)$ at $x = c_1$, c_3 the value of c_2 at $x = c_2$, and so on, then $|g(x)| < \frac{1}{2}|g(c_1)| < \frac{1}{2^2}|g(c_2)| < \dots < \frac{1}{2^n}g(c_n)$, for any n; but $|g(x)|$ is continuous, and so bounded in (ξ_r, ξ_{r+1}), and $\frac{1}{2^n} \to 0$, and therefore $|g(x)| = 0$ at all points of (ξ_r, ξ_{r+1}), i.e. $\omega(x) = \psi(x)$ in (ξ_r, ξ_{r+1}). In particular $\omega(x) = \psi(x)$ at $x = \xi_r$ and at $x = \xi_{r+1}$, and so, by the foregoing argument, $\omega(x) = \psi(x)$ throughout both (ξ_{r-1}, ξ_r) and (ξ_{r+1}, ξ_{r+2}). Thus step by step the equality extends to all the intervals $(\xi_\mu, \xi_{\mu+1})$, and so to the whole interval (x_0^*, x_1^*). We have proved that if $\omega(x) = \psi(x)$ at one point in (x_0^*, x_1^*) then the equality holds throughout the interval; but

$$\omega(a) = \psi(a) = b,$$

and therefore $\omega(x) = \psi(x)$ throughout (x_0^*, x_1^*).

15.94. Let R be the rectangle $(a,b)(c,d)$; since $u(x,y)$ is continuous in $(a,b)(c,d)$ it is x-continuous in (a,b) for any y in (c,d).

Hence by Example 13.91, if (p,Q) and (P,Q) lie in $(a,b)(c,d)$, there is a p^* between p and P such that

$$\frac{f(P,Q)-f(p,Q)}{P-p} = u(p^*, Q).$$

Similarly

$$\frac{f(p,Q)-f(p,q)}{Q-q} = v(p, q^*),$$

therefore

$$f(P,Q)-f(p,q) = (P-p)u(p^*,Q)+(Q-q)v(p,q^*)$$

$$= (P-p)\{u(p,q)+o(n)\}+(Q-q)\{v(p,q)+o(n)\},$$

provided $P-p = O(\lambda_n)$, $Q-q = O(\lambda_n)$, since $u(x,y)$ and $v(x,y)$ are continuous in R.

Hence $f(x,y)$ is differentiable in R with x-derivative $u(x,y)$ and y-derivative $v(x,y)$.

15.95. Let $\qquad \phi(x,y,k) = f(x,y+k) - f(x,y),$

then

$$f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)$$
$$= \phi(x+h,y,k)-\phi(x,y,k)$$
$$= h\phi_x(x+\theta_1 h,y,k) \quad \text{by the mean-value theorem,}$$
$$= h\{f_x(x+\theta_1 h,y+k)-f_x(x+\theta_1 h,y)\} = hk\{f_{xy}(x+\theta_1 h,y+\theta_2 k)\},$$
$$\text{by the mean-value theorem,}$$
$$= hk\{f_{xy}(x,y)+0(n)\}, \quad \text{provided } h = 0(p_n),\ k = 0(p_n),$$

since f_{xy} is continuous.

But

$$\frac{\phi(x+h,y,k)}{k} = \frac{f(x+h,y+k)-f(x+h,y)}{k} \to f_y(x+h,y), \quad \text{as } k \to 0,$$

and
$$\frac{\phi(x,y,k)}{k} = \frac{f(x,y+k)-f(x,y)}{k} \to f_y(x,y), \quad \text{as } k \to 0.$$

Hence

$$\frac{f_y(x+h,y)-f_y(x,y)}{h} = f_{xy}(x,y)+0(n-1), \quad \text{provided } h = 0(p_n),$$

which proves that $f_{yx}(x,y)$ exists and equals $f_{xy}(x,y)$.

15.96. For non-zero values of x and y,

$$f_x = 4\frac{x^3}{y}\sin\frac{y^2}{x}-x^2y\cos\frac{y^2}{x}, \qquad f_y = -\frac{x^4}{y^2}\sin\frac{y^2}{x}+2x^3\cos\frac{y^2}{x},$$

$$f_{xy} = -4\frac{x^3}{y^2}\sin\frac{y^2}{x}+7x^2\cos\frac{y^2}{x}+2xy^2\sin\frac{y^2}{x} = f_{yx}.$$

Furthermore

$$f_x(x,0) = \lim_{X\to x}\frac{f(X,0)-f(x,0)}{X-x} = 0, \quad \text{for all } x.$$

$$f_x(0,y) = \lim_{x\to 0}\frac{f(x,y)-f(0,y)}{x} = \lim_{x\to 0}\frac{x^3}{y}\sin\frac{y^2}{x} = 0, \quad y \neq 0,$$

since $\left|\sin\frac{y^2}{x}\right| < 1$, etc.

XVI

16. Let $f(x,y)$ take a maximum or minimum value c at (a,b) on $g(x,y) = 0$; the tangent to $g(x,y) = 0$ at (a,b) is $(x-a)g_a+(y-b)g_b = 0$ and the tangent to $f(x,y) = c$ is $(x-a)f_a+(y-b)f_b = 0$. But $f+\lambda g$ is stationary at (a,b), whence $f_a+\lambda g_a = 0 = f_b+\lambda g_b$ and so

$$(x-a)f_a+(y-b)f_b = -\lambda\{(x-a)g_a+(y-b)g_b\}.$$

16.2. Use Example 16.

16.3. $x = 6$, $\cos\theta = \frac{1}{2}$.

16.4. Minimum values $4/(4+2\sqrt{2})$, $4/(4-2\sqrt{2})$.

16.6. Consider $XY(X+Y)^3$, where $X = x^3$, $Y = y^3$.

16.92. Let $x = -f$ be the directrix and $(a\cos\theta, a\sin\theta)$ the focus; equation of the parabola is

$$y^2 - 2ay\sin\theta - 2x(f + a\cos\theta) + a^2 - f^2 = 0.$$

16.93. Envelope is $x = x\{\alpha(t), t\}$, $y = y\{\alpha(t), t\}$, where the function $\alpha(t)$ is to be so chosen that this curve touches each of $x = x(\alpha, t)$, $y = y(\alpha, t)$. Condition for contact is

$$\frac{\partial x}{\partial t}\left(\frac{\partial y}{\partial \alpha}\frac{d\alpha}{dt} + \frac{\partial y}{\partial t}\right) = \frac{\partial y}{\partial t}\left(\frac{\partial x}{\partial \alpha}\frac{d\alpha}{dt} + \frac{\partial x}{\partial t}\right),$$

that is,
$$\frac{d\alpha}{dt}\cdot\frac{\partial(x,y)}{\partial(\alpha,t)} = 0,$$

and so
$$\frac{\partial(x,y)}{\partial(\alpha,t)} = 0, \quad \text{since } \frac{d\alpha}{dt} \neq 0.$$

16.94. Regard a, b as functions of a single variable t, satisfying $\lambda(a,b) = 0$. Envelope satisfies $\phi = 0$, $\partial\phi/\partial t = 0$. But

$$\frac{\partial\phi}{\partial t} = \frac{\partial\phi}{\partial a}\cdot\frac{da}{dt} + \frac{\partial\phi}{\partial b}\cdot\frac{db}{dt} \quad \text{and} \quad \frac{\partial\lambda}{\partial a}\frac{da}{dt} + \frac{\partial\lambda}{\partial b}\frac{db}{dt} = 0,$$

whence $\dfrac{\partial(\phi,\lambda)}{\partial(a,b)} = 0$; conversely if $\dfrac{\partial(\phi,\lambda)}{\partial(a,b)} = 0$ and $\dfrac{\partial\lambda}{\partial a}\dfrac{da}{dt} + \dfrac{\partial\lambda}{\partial b}\dfrac{db}{dt} = 0$ then

$$\frac{\partial\lambda}{\partial b}\cdot\frac{\partial\phi}{\partial t} = 0 \quad \text{and so} \quad \frac{\partial\phi}{\partial t} = 0, \quad \text{since } \frac{\partial\lambda}{\partial b} \neq 0.$$

Normal at (a,b) on $\psi(x,y) = 0$ is given by $(x-a)\psi_b - (y-b)\psi_a = 0$, $\psi(a,b) = 0$, whence envelope satisfies

$$(x-a)(\psi_{ab}\psi_b - \psi_{bb}\psi_a) - (y-b)(\psi_{aa}\psi_b - \psi_{ab}\psi_a) = \psi_a^2 + \psi_b^2.$$

Solving this equation with that of the normal we obtain the evolute as given in Example 15.41.

16.95. Family of rectangular hyperbolas is

$$x^2 + xy\tan\theta - y^2 - x\sec\theta + 2y\sin\theta - \sin^2\theta = 0.$$

16.96. There is a maximum at $x = t$, $y = t$, where

$$mAt^{-m-1}(1+2^{-m-1}) = nBt^{-n-1}(1+2^{-n-1}).$$

XVII

17.1. Divide the triangle T bounded by $y = x$, $y = 4-x$, and $x = 1$ into T_1 bounded by $y = x$, $y = 2$, $x = 1$ and T_2 bounded by $y = 4-x$, $y = 2$, $x = 1$, and apply $\int\limits_T = \int\limits_{T_1} + \int\limits_{T_2}$.

17.2.

$$\int\limits_C f(x^2+y^2)\,dxdy = \lim_{n\to\infty} \int\limits_{1/n}^{1} \int\limits_0^{2\pi} f(r^2)\frac{\partial(x,y)}{\partial(r,\theta)}\,drd\theta, \quad x = r\cos\theta,\ y = r\sin\theta,$$

$$= \lim_{n\to\infty} \pi \int\limits_{1/n}^{1} 2rf(r^2)\,dr = \lim_{n\to\infty} \pi\{F(1) - F(1/n^2)\} = \pi F(1).$$

17.4. If C is the boundary of the area required then

$$\int_C dxdy = \int_{q_2}^{q_1}\int_{p_2}^{p_1} \frac{\partial(x,y)}{\partial(u,v)}\,dudv$$

where $u = x^2/y$, $v = y^2/x$, so that $xy = uv$ and therefore $x = u^{\frac{2}{3}}v^{\frac{1}{3}}$, $y = u^{\frac{1}{3}}v^{\frac{2}{3}}$, and $\dfrac{\partial(x,y)}{\partial(u,v)} = \frac{1}{3}$. Accordingly

$$\int_C dxdy = \tfrac{1}{3}(p_1-p_2)(q_1-q_2).$$

17.5. Express the double integral in two ways as a repeated integral.

17.6. Under the given transformation the ellipse $\dfrac{x^2}{c^2\,\mathrm{ch}^2u}+\dfrac{y^2}{c^2\,\mathrm{sh}^2u} = 1$ becomes the line segment $0 \leqslant v < 2\pi$, $u =$ constant, and so R transforms into the rectangle $0 \leqslant v < 2\pi$, $0 < u < \mathrm{sh}^{-1}b/c$.

$$1/r_1+1/r_2 = 2c\,\mathrm{ch}\,u/(c^2\,\mathrm{ch}^2u-c^2\cos^2v)$$

and the Jacobian of the transformation is $c^2(\mathrm{ch}^2u-\cos^2v)$, whence

$$\int_R (1/r_1+1/r_2)\,dxdy = \int_0^{2\pi}\int_0^{\mathrm{sh}^{-1}b/c} 2c\,\mathrm{ch}\,u\,dudv = 4\pi b.$$

17.8. Take $0 < \epsilon < \sqrt{X}$, and let Q be the quadrant of the circle $x^2+y^2 = 2X$ which contains the square $(0,\sqrt{X})(0,\sqrt{X})$, and let Q^* be Q less the quadrant of the circle $x^2+y^2 = \epsilon^2$. Then

$$\left(\int_\epsilon^{\sqrt{X}} t^{2n+1}e^{-t^2}\,dt\right)^2 = \left(\int_\epsilon^{\sqrt{X}} t^{2n+1}e^{-t^2}\,dt\right)\left(\int_\epsilon^{\sqrt{X}} u^{2n+1}e^{-u^2}\,du\right) = \int_\epsilon^{\sqrt{X}}\int_\epsilon^{\sqrt{X}} (ut)^{2n+1}e^{-u^2-t^2}\,dudt$$

$$< \int_{Q^*} (ut)^{2n+1}e^{-u^2-t^2}\,dudt = \left(\int_\epsilon^{\sqrt{(2X)}} r^{4n+2}e^{-r^2}r\,dr\right)\left(\int_0^{\frac{1}{2}\pi} \sin^{2n+1}\theta\cos^{2n+1}\theta\,d\theta\right),$$

$t = r\cos\theta$ and $u = r\sin\theta$,

$$= \frac{(n!)^2}{4\{(2n+1)!\}}\int_{\epsilon^2}^{2X} x^{2n+1}e^{-x}\,dx, \quad \text{where } x = r^2,$$

and so, letting $\epsilon \to 0$,

$$\left(\int_0^{\sqrt{X}} t^{2n+1}e^{-t^2}\,dt\right)^2 < \frac{(n!)^2}{4\{(2n+1)!\}}\int_0^{2X} x^{2n+1}e^{-x}\,dx < \frac{(n!)^2}{4\{(2n+1)!\}}\int_0^{2X+1} x^{2n+1}e^{-x}\,dx;$$

but $$\int_0^X x^ne^{-x}\,dx = 2\int_0^{\sqrt{X}} t^{2n+1}e^{-t^2}\,dt, \quad x = t^2,$$

whence taking $X = n$ we find $u_n^2 < u_{2n+1}$.

INDEX